AD_____

Award Number: W81XWH-05-1-0101

TITLE: Development and Evaluation of Sterographic Display for Lung Cancer Screening

PRINCIPAL INVESTIGATOR: Xiao Hui Wang, M.D., Ph.D.

CONTRACTING ORGANIZATION: University of Pittsburgh
Pittsburgh, PA 15260

REPORT DATE: December 2008

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1 Dec 2008	Final	8 Nov 2004 – 7 Nov 2008

4. TITLE AND SUBTITLE

Development and Evaluation of Sterographic Display for Lung Cancer Screening

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-05-1-0101

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Wang, Xiao Hui, Good, Walter F, Fuhrman, Carl R, Rockett, Howard E, Gur, David

E-Mail: xwang@mail.magee.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Pittsburgh
Pittsburgh, PA 15260

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

See next page.

15. SUBJECT TERMS
lung cancer screening, stereo display, volumetric rendering, observer performance study

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	58	USAMRMC

ABSTRACT

The main purpose of this project is to investigate the feasibility and efficacy of using a stereo display workstation for lung cancer screening on CT images. The tasks included in this project are development and evaluation of stereo image projection and display for chest CT images, observer performance evaluation for the stereo display, and stereo feature analysis and comparison to the conventionally used display methods for lung cancer detection. During the funding period, we have made progress in following tasks. 1. We have built stereo display workstation for chest CT images and investigated effects of several commonly used compositing methods for nodule representation and detection in stereo CT images. Among these methods, conventional maximum intensity projection (MIP) produced the highest image contrast, but gave ambiguities in local geometric detail and texture, whereas averaging compositing resulted in the lowest contrast, but preserved geometric details. Distance-weighted MIP partially recovered geometric information, which was lost in images composited by conventional MIP, therefore is the best compositing method for stereo display. 2. Consensus truth of the cases collected for this project has been done bythree experienced radiologists. 3. A pilot observer performance study was conducted. Six radiologists participated the pilot observer performance study. The study has three display modes, conventional slice-by-slice mode, conventional MIP display mode and stereo display mode. The performance of lung nodule detection are examined and compared for the three modes with Free-response Receiver Operating Characteristic (FROC) statistic method. The results indicate that the stereo display achieved the best performance followed by the slice-by-slice display, and the conventional MIP display gave the worst performance, although there is no statistically significant difference between the three display modes. Subjective assessment indicates that the stereo display was well accepted by the radiologists. 4. We have explored advanced features for the stereo display. We tested the feasibility and efficacy of performing 3D rendering on GPUs (Graphics Processing Units) for stereo display of medical images. Our GPU-based program achieved real-time rendering, real-time displaying and real-time interactive controls by radiologists, which is desirable and necessary for prompt and accurate medical diagnosis. We also investigated spectrophotometric characteristics of stereographic image pairs to further understand the characteristics of stereo imaging and displaying. We analyzed differences in spectrophotometric characteristics between images acquired during stereographic imaging, and found that though uniform global differences can easily be corrected by applying traditional histogram matching techniques, these methods are not capable of dealing with differences that are object or distance dependent. We have developed a procedure to locally adjust visual characteristics of one image in a stereo pair to match the alternate image. The fullyautomatic procedure is able to remove visible differences in most cases, therefore enhance the quality of stereo 3D visualization. 5. A main observer performance study was conducted, which used larger database, more readers, and improved study design based on the feedbacks from the pilot study. Averaged areas under ROC curves (Az) of the performance on stereo, MIP and slice-by-slice displays are 0.67±0.06, 0.65±0.06, 0.65±0.04, respectively. Our study results indicate that 3D representation and visualization can be further improved in terms of rendering scheme, flexibility of volume thickness and cutting planes, and real-time image processing for better comprehension and easy maneuver of 3D images.

Table of Contents

INTRODUCTION

Lung cancer is a leading cause of death in the United States [1,2]. The results from several large lung cancer screening studies indicate that early detection and treatment can reduce mortality rate in most types of lung cancer cases [3-6]. Currently, low-dose CT scanner is a primary tool used for lung cancer screening. For each screening case, a set of image slices covering entire lung area is generated and viewed on display workstations. Despite of 3D format of CT datasets, the conventional reading method for lung CT image interpretation is to read images slice-by-slice. This reading method requires radiologists to mentally reconstruct images in 3D space from a set of 2D images to differentiate normal tubular structures from nodules. Furthermore, with improved technology for CT scanner, higher resolution imaging techniques produce more images per scan, which eventually will exceed radiologists' ability to read cases in slice-by-slice mode. The need of 3D data presentation of CT images has become crucial for ever-increasing numbers of images generated from CT scanner and for improvement of radiologists' performance on image data interpretation. We have proposed to develop a stereo display workstation for reading lung CT images. Stereopsis is the mechanism used in human vision system to perceive objects in our three dimensional space. The 3D display using stereoscopic projection should produce a natural and efficient solution for 3D data presentation. In this proposal, we hypothesized that the efficacy of lung cancer screening using CT scanned images can be increased by use of a suitable designed stereoscopic display. Specifically, we expect that both efficiency, and accuracy for the detection of lung nodules, will be increased significantly over what can be achieved when reading cases in currently used display modes. To achieve the goals in this proposal, we have specified our aims as followings:

1) Develop and integrate the hardware and software required to implement a stereoscopic display tailored to chest CT images.
2) Use a subset of lung cancer cases, verified either by pathology or by followup, to evaluate the display system.
3) Perform a retrospective study to measure relative accuracy and reading efficiency, for detection and classification of lung nodules, between three display modes including stereoscopic 3D mode from this project, and other two commonly used modes, slice-by-slice and maximum-intensity-projection (MIP) thick slice.

BODY

Integrate Hardware (task 1)

To implement the stereo display workstation, certain requirements of hardware components and hardware constructions need to be fulfilled before integration of the hardware and software for viewing stereographs.

The integrated hardware for stereo display consisted of a PC computer equipped with stereographic card, a programmable keypad, a monitor, a signal synchronizer, and a shutter glasses.

Computer – A 2.8 GHz AMD Athlon 64 personal computer was configured for stereo image processing and stereographic display. The computer has 3 hard disk units connected via RAID technology to create maximal disk capacity of 400 GH that can support massive data computing and real-time display for stereo chest CT workstation. The NVIDIA Quadro FX1100 graphic card installed in the computer is stereo capable that provides display buffers and OpenGL/DirectX support necessary for stereographic presentation. The computer's performance and stereographic capability was tested for the optimal level to the tasks in this project.

Keyboard – Several functionalities for controlling stereo display window were encoded into a programmable keypad. The key-controlled features include, but not limited to, adjusting window/level settings, scrolling sequential images in a case, changing number of image slices that are used for composing stereo images, and toggling markers used for detected nodule. In lieu of regular keyboard, the keypad we reconstructed is tempered

to the needs of lung nodule detection and classification in the stereo display. By simply pushing on the functional keys, radiologists can manage the display to the level of optimal viewing and task specific.

Stereo output ─ Three parts, monitor, signal emitter and shutter glasses, in the integrated hardware are mainly responsible for stereo image output. In order to perceive stereoscopic view of the images, left and right images need to be viewed separately by each of corresponding eyes at frame rate of at least 60 HZ for each eye. The graphic card was configured to alternate displayed left and right images and output the signal through the emitter to synchronize the shutter glasses to the refresh cycle of the monitor. The monitor with refresh rate of 144 HZ is used to give flicker-free stereo images viewed through shutter glasses.

Develop 3D Geometric Projection Algorithms and Transparency-Contrast Models (tasks 2, 3)

Display a 3D chest CT dataset in stereo requires that two 2D projections, corresponding to left and right eye views (stereo pair), be performed. Projection methods consist of geometric projection model and illumination model. The interaction between geometric projection model, illumination model and optical characteristics at each voxel can be integrated along rays to calculate the values of pixels on 2D projection. For CT images, it is customary, for display, to assign brightness values having a monotonic, but nonlinear relationship to X-ray attenuation coefficients, as measured in Hounsfield units. In keeping with this, we assumed that each CT voxel has a neutral color (some shade of gray) and a brightness value that is some affine mapping of its Hounsfield value, with pixels corresponding to low X-ray attenuation appearing as darker intensities and pixels corresponding to higher X-ray attenuation appearing brighter.

In stereopsis, contrast is important for depth perception [7-8], and subjective evaluations from our studies indicate that optimization of local contrast is necessity for detecting fine structures in stereo displays of chest CT data. Monocular occlusions from the perspective transformation and from averaging effects of transparency models incorporated in traditional volumetric reconstruction algorithms possibly reduces contrast in the reconstructed stereo pairs with increased image depth from stacks of CT slices. Our goal, therefore, was to maximize the detectability of nodules, but at the same time, maintain sufficient geometric fidelity to allow detected nodules to be accurately characterized. The issues of clearly displaying other structures of interest such as vessels, bronchi and bone and incorporation of monoscopic depth cures were also addressed and implemented in the study.

a. Geometric Models

The two geometric projection models that are relevant in the context of this application are the orthographic and the perspective transformations. Orthographic projection, which is used by traditional Maximum Intensity Projection (MIP), is not compatible with realistic stereographic projection. Perspective transformations have been used to a limited extent in the medical environment [9], but their added complexity is not always justified for monoscopic viewing. Geometric perspective is one of the more important monoscopic depth cues when familiar scenes are being viewed, though the extent, to which it would be of value for viewing the interior of lungs, is not known.

To test the prospective use of perspective transformation for stereo view of lung images, we applied perspective transformation in stereo image compositing algorithm to a set of chest CT images. We assumed that the topmost slice being displayed is the same distance as the screen of the physical display, the viewing distance between a viewer and the screen is 45-cm, and the interocular distance is 6.5-cm. We have maintained these conditions throughout this work, but in our studies we found that stereoscopic convergence is not appreciably affected by the exact seating position ─ as long as the eyes are maintained level relative to the monitor.

It is generally aware that volumetric lung images acquired by current CT systems have different resolutions in the axial direction than in planes perpendicular to that direction. The axial resolution is more a function of the user's specification to the reconstruction algorithm, than to the acquisition protocol. The anisotropic nature of the dataset (e.g., voxels are typically 0.7-mm × 0.7-mm × 2.5-mm), and our choice of a non-orthogonal projection, necessitates resampling the data to obtain the actual values used in ray casting. The most common

method used for interpolation in computer graphics is trilinear interpolation because it is easy to implement, computational efficient and comparable to other more sophisticated interpolation methods in terms of visual effect. We used trilinear interpolation to resample the data and insert effectively the virtual slices between each pair of real slices to achieve final pixel dimension close to isotropic.

The results from MIP volumetric rendering with perspective transformation produced stereo images that had true 3D depth and spatial differentiation of interior lung structures. Comparing visually to perspective transformation, MIP rendering with orthogonal transformation generated only monoscopic 3D view and no information of depth or of the geometrical relationships between structures.

b. Implement monoscopic depth cues

Disparity between stereoscopic views is the primary depth cue being studied in this work. However, there exist monoscopic depth cues that greatly assist viewers in achieving stereopsis, and it is desirable to include these cues, to the extent they are not in conflict with the goal of optimizing radiographic interpretation. Two monoscopic cues were incorporated in the stereo compositing methods, geometric depth cues and brightness cue.

Geometric cues ─ The main geometric depth cue we have included was geometric perspective which, as was discussed above, was the transformation adopted exclusively for rendering stereoscopic images in this project. A second geometric cue we have included was the occlusion of voxels by intervening structures to determine the amount of illumination projected on display plane for an object.

Brightness cue -- We have modeled brightness variation with respect to depth change. To model a systematic reduction in brightness values with increasing depth or distance, we assumed that each slice has a fixed optical density that reduces the brightness of slices lying behind it. In the actual implementation, each slice was assumed to have the same fixed optical density and weighing factors, derived from a geometric sequence, was applied so as to achieve a ratio of weights, K, between the back slice and the front slice. For N slices, and assuming the total of all weights equals 1, the weights W_i, $i=0, \ldots, N-1$, was calculated as follows:

$$W_i = \frac{1 - K^{\frac{1}{N-1}}}{1 - K^{\frac{N}{N-1}}} K^{\frac{i}{N-1}}$$

In the case of our distance-weighted MIP projections, the front slice weight was 1, and the sum was irrelevant. In our preliminary studies we found that K=0.5 provides a reasonable indication of depth without excessively reducing the brightness of the deeper slices.

c. Illumination Models

Volume rendering, which was adopted for our stereo image compositing, attempts to identify and classify all voxels of an object, and to assign optical properties (e.g., transparency, color and brightness) to each voxel and is particularly useful in applications, such as chest CT, where the structures of interest are rather sparsely distributed.

The challenging task in volume rendering is to develop an optimal brightness / contrast / transparency model for assigning optical characteristics to each rendered voxel. Because the intrinsic property being imaged with CT is actually X-ray attenuation coefficient, the assignment of optical properties to CT voxels is entirely artificial and should be done in a manner that best facilitates the interpretation of the image data. Two illumination models that are suitable for chest CT data have been examined and compared.

1. Averaging model ─ this compositing method that is commonly applied in volume rendering averages voxel values along rays and effectively reduces the contrast of smaller objects, which may appear in only a few slices,

as the displayed volume is increased by adding slices. Nevertheless, volume rendering with averaging does preserve spatial information such as textures and local geometry.

For chest CT images, instead of simply averaging voxel values along rays, we adopted a light emission / transmission / occlusion model that assumed that each voxel emits light in proportion to its brightness when CT images are displayed at a normal window and level for nodule detection and that uses distance information (distance-weighting factors) to determine the amount of this emitted light that reaches the projection plane (appendix A). Specifically, it was assumed that each slice has a fixed optical density that decreases the brightness of slices lying behind it. The total of all distance weights was equal to one. The ratio of weights between the last (the slice with the largest distance from screen) and first slice (the slice at screen level) controls the level of transparency for a given volume.

We studied a range of these ratios for lung CT images and empirically set the ratio to 0.5 to achieve a balance between the use of brightness weighting as a depth cue and visibility of the back slice. The final value for a voxel is the sum of distance-weighted pixel values in a perspective transformation ray.

2. MIP models – by using the maximum pixel intensity along each ray of the projection, MIP is designed to maximize contrast in situations where sparsely distributed objects are being viewed against a dark background, which is the situation that occurs in projecting thick volumes of the lung. Its main deficiency is that it does not preserve local spatial structure such as texture and local geometry. This is because a bright pixel appearing in a view for one eye will not necessarily have a corresponding bright pixel that appears in the view for the opposite eye.

In our stereographic compositing with MIP principle, we have experimented two different ways of applying MIP compositing. First, we used a perspective projection in which the maximum value along each ray was used as the projected value. As we mentioned above, this approach generally is not possible for an observer's vision system to unambiguously match corresponding points between the two views because the projected voxels may be different between the stereo views. In practice, we found that the ambiguity in matching corresponding points between views primarily affects fine detail and does not interfere with the detection of objects composed of large clusters of bright voxels, although the exact shapes of these objects cannot be determined unambiguously.

In an attempt to incorporate a geometric cue common in traditional stereo projection methods, but also preserve the contrast advantage of MIP, we in turn used a perspective projection in which the maximum along each ray has been weighted, based on distance. The maximum (nearest) and minimum (farthest) weights were determined empirically, and weights for slices between the first and last were calculated by suing a geometric sequence, assuming a fixed optical density for each slice.

d. Compare compositing methods for lung nodule detection and characterization on stereo CT images

We compared averaging, MIP and distance-weighted MIP applied to various lung nodule types, sizes, and locations. MIP and distance-weighted MIP produced higher local contrast than compositing by averaging. Unlike averaging methods, which sacrifice contrast to take account of each voxel in a volume, the conventional MIP method is able to retain contrast in cases in which the object being viewed includes voxels that are brighter than the superimposed tissue. Applied to the task of lung nodule detection, despite a lack of geometric fidelity, conventional MIP images generally produce high local contrast that separates a nodule from its background and therefore enhances detection performance.

However, improvements in nodule visibility with the MIP method do not apply in certain cases in lung CT images. For a nodule to be detected with conventional MIP, it must contain some voxels that are brighter than its background. The case of a nodule overlying a rib is a particular concern, although it occurs relatively infrequently in projections of axial slabs unless the nodule is very close to a rib or the slab being viewed is

relatively thick. Most voxels in such a nodule will not be as bright as voxels in the rib, and the nodule may be almost invisible.

One benefit of the distance-weighted MIP projection is that it can decrease the brightness of the background. As the position of a slab is changed, relative brightness weighting factors between voxels at different axial positions will change, and, in many instances, there will be an axial position at which a nodule will appear brighter than a rib, which can increase the likelihood that an obscured nodule becomes visible. Our results indicate that distance-weighted MIP partially recovered geometric information lost in conventional MIP by incorporating a distance cue into the compositing, and at the meantime, distance-weighted MIP had image contrast nearly equivalent to that produced by the conventional MIP method.

In addition to detection, nodule characteristics are essential and critical for clinically differentiating benign for malignant. Although conventional MIP is superior for detection, it was outperformed by the averaging method in terms of characterization. When comparing the three compositing methods visually for nodule characterization, the spiculated nodule border is clearly visible in stereo pairs composited with the averaging method, whereas this characteristic is not preserved in some of the stereo pairs composited with the conventional MIP method, especially those composited from thicker slabs. Distance-weighted MIP partially overcomes the problem of conventional MIP, and speculated borders are still visible in the thicker slabs. For the smooth borer of a nodule, we observed a similar phenomenon. The geometric relationship is well presented with gradient changes in intensities along the smooth border of the nodule in the stereo pair composted with the averaging method. The same nodule is shown with lack of geometric fidelity in the stereo pair composited with conventional MIP. The lack of fidelity in module shape and geometric representation in conventional MIP images are attributed to the nature of MIP compositing, in which the two views may be based on projections of different voxels. Conversely, the averaging method is shown in this study to faithfully retain the characteristics of the nodules, including structural, spatial, and geometric information.

e. Test combinations of MIP and averaging

Each compositing method as described above has a particular balance of advantages and disadvantages, and may be optimal in certain situations but inappropriate for others. Based on our preliminary results, we believe that two rendering modes will need to be available to viewers -- one optimized for nodule detection and one optimized for nodule characterization. Thus, in practice, it may actually be advantageous to view a volume with a range of compositing methods.

Because MIP appears to be best for nodule detection, and some form of voxel averaging is best for charactering detected nodule, we have implemented and used two separate rendering methods, i.e. distance-weighted MIP and averaging, in a single stereoscopic display mode. The intention is that after a nodule has been detected in the distance-weighted MIP, the thickness of the displayed volume can be adjusted to include only slices that contain the nodule, and this volume can be displayed by using an averaging compositing method so the nodule can be characterized more accurately.

Write Display Software (task 4)

The display software provided for the stereo viewing has been developed. The functionalities built into the software enabled readers to have real-time interaction with slice navigation through entire lung sections, to change viewing thickness of lung sections, and to adjust brightness / contrast of displayed images (Appendix B).

a. Write user interface

The display software was written as a Windows application implemented specifically with Win32 including MFC (Microsoft Foundation Classes), and SGI (Silicon Graphics Inc.) OpenGL (Open Graphics Library) language. The Windows API provided Windows framework and functional utilities of the user interface. The OpenGL functions enabled stereographic display by manipulating display functions on graphic processing unit. A typical user interface is shown in figure 1.

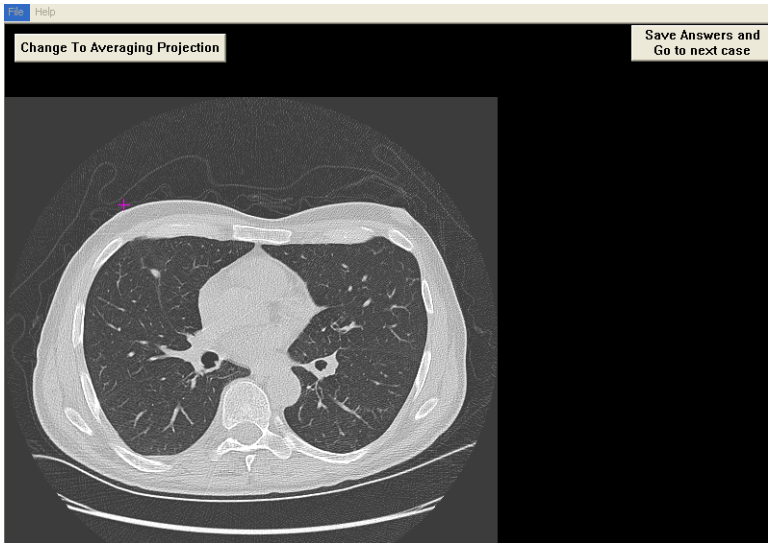


Figure 1. Window-based user interface.

b. Write routines for prestaging projections

In order to have users to control the thickness (i.e., number of slices in a slab) and the axial position of stacks of 2D CT slices, all projections were precalculated and stored on the display's hard disk. For each case, the files were organized into a two-dimensional linked list that allows projections to be accessed by thickness and axial position in real-time. Software was written to render stereo image pairs from the CT data by the stereo projection algorithms described in tasks 2, 3.

For the comparisons, the software also included rendering traditional monoscopic MIP images. To be consistent with commercially available CT displays, monoscopic MIP images were generated using a standard orthogonal projection. The slice-by-slice display mode required only minimal additional software because, by design, this mode is actually a limiting case (i.e., volume thickness = 1 slice) of both the MIP and stereo modes.

All computer code was written in C++ language and tested thoroughly prior to its use in this study.

c. Write routines for window/level control

Window / level (contrast / brightness) mechanism was built into the display workstation with a default setting of 2000 / -500-HU for lung CT images. Real-time adjustment of window / level was enabled through encoded keys on the keypad.

d. Write routines for control of slice thickness and axial position

To achieve real-time performance in changing axial position in the CT volume and in changing the thickness of the displayed volume, stereo pairs corresponding to all admissible combinations were precalculated, organized into a 2D linked list, and stored on the hard drive. This required approximately 1-GB of disk storage per case, but was able to achieve nearly instantaneous response to changes in axial position or slice thickness.

Implement Secondary Features (Task 5)

We continued development of certain display features we consider to be of value in this application. These features are important in broadening the applicability of stereo displays for chest CT. The main secondary features implemented were volume projection and rotation in real-time with GPU (Graphics Processing Unit) card.

Recent advanced commodity GPUs are very efficient at manipulating and displaying computer graphics for a range of complex algorithms. The advanced features of GPUs are especially useful to medical practice, in which

data interpretation is timely dependent, extensive interactions are required, and multiple format of data presentation in real time is desired for different diagnostic purposes.

Applying programmable GPUs is likely a solution for real-time stereo image compositing and display. In this particular application, the tasks that the GPUs can facilitate for lung CT stereo display include stereo pair compositing from CT data set at desired viewing position and viewing volume, multiple rendering algorithms, brightness/contrast adjustment, and image rotation.

The GPU-based program has achieved real-time rendering and real-time display effect without any perceptive delay in each successive frame rendering and display following a user controlled frame switch command. We found no difference in frame rate between MIP and average renderings. To test and demonstrate the real-time stereo rendering process on GPU card, we used lung CT images for quantitative measurements of the performance shown in table 1 and 2.

Table 1. Frame rates measured as stereo pairs per second for rendering on GPU card and CPU (Central Processing Unit) card at different number of interpolated slices.

Number of interpolated slices	GPU (stereo pairs per second)	CPU (stereo pairs per second)
1	103.3	-
9	20.1	1.3
15	13.2	0.8
21	10.1	0.5
33	6.6	0.3
45	5.0	0.2

Table 2. Frame rates measured as stereo pairs per second for rendering on GPU card with and without rotation implementation.

Number of interpolated slices	Rotation implemented (stereo pairs per second)	Without rotation (stereo pairs per second)
1	103.3	103.3
3	44.4	44.4
5	33.7	33.7
9	20.1	20.1
15	13.2	13.2
21	10.1	10.1
33	6.6	6.6
45	5.0	5.0

Our results indicate that programming on GPUs can not only avoid lengthy process of precalculation and overloaded disc space, but also provide some functionalities that would be virtually impossible for the prestaged process, such as rotations and real-time interpolations (for example, changing image resolution). The GPUs solution has shown to be efficient for real-time stereo pair renderings and display. We have submitted our results for peer-reviewed publication (Appendix A).

To further understand stereo image features and stereo display characteristics, we investigated spectrophotometric characterstics of stereographics of stereographic image pairs. Differences in spectrophotometric characteristics between images acquired during stereographic imaging may significantly reduce the effectiveness of their subsequent display or analysis. While uniform global differences can easily be corrected by applying traditional histogram matching techniques, these methods are not capable of dealing with differences that are object or distance dependent. We have developed a procedure to adjust locally, visual characteristics of one image in a stereo pair to match the alternate image. Objects, and their boundaries, are

segmented in both images. Non-uniform regions and very small objects are either suppressed or combined into a single large region, while larger objects are retained. Local pattern matching, by varying the horizontal displacement between images, allows a correspondence to be established between boundaries of the objects on the two images, and hence a correspondence between connected components. For each pair of corresponding connected components, a linear correction function that minimizes the sum-of-squares difference is determined. Each pixel in the image to be corrected is adjusted by interpolating between all of the correction functions based on the distance of the pixel from each of the centers-of-mass of the individual connected regions. The fully automatic procedure is able to remove visible differences in most cases.

Write Software Needed To Perform Study (task 6)

In the main study as well as the pilot study, readers are required to identify possible nodules and then characterize the detected nodules in three display modes. A software was written to implement functions required for the study, which include management of study cases and display modes, case randomization, electronic scoring of detected nodules, scaled onscreen ruler for nodule measurement, markers for detected nodules, and data archive. All implementations were written in C++ language and have been tested rigorously to meet the specifications of the study.

a. Write 3D cursor routine for making nodule locations

We were able to make a stereo cursor corresponding to the perspective view of displayed volume. In specific, the cursor can be specifically placed at a particular location in the 3D volume space. The size of the cursor is properly transformed to reflect the size perceived at a distance based on the spatial location of the cursor in the volume. With the stereo cursor, a reader can mark a nodule location (with 3 axes of x, y, and z) in the volume. At the present, we have placed the cursor in the middle of the viewing slab. To make a navigable stereo cursor, we just need to encode mouse wheel with the stereo cursor routines, since all the mechanisms have been laid out.

Besides serving as a location indicator, the stereo cursor was also made as an onscreen measure for nodule size. Horizontal or vertical line of the cursor represents unit of length that can be used for size approximation when the cursor is placed over a structure.

b. Implement case randomization

In observer performance study, case randomization is a procedure to assure that a reader's response is not biased by the order and the time of the cases presented. Our randomization routines were designed to generate random case list, remove the cases that have been examined and then present the cases for each reader at each reading session. Because of the multiple modes we need to test in this study, the randomization routines have also implemented function that checks status of a case across the modes.

c. Record findings

Electronic scoring form was implemented in the study software. When a reader clicks on a possible nodule, a scoring form with study questions will show on the screen. The questionnaire can be answered in the form of either check boxes or radio buttons. The answers as well as nodule's location will be saved after reader finishes scoring or will be neglected if reader chooses to delete the answers. A detected nodule is toggled with a maker and on/off states of the marker can be controlled at reader's will.

d. Write routines for monitoring readers' reading patterns

It was our intention to incorporate routines into the study software for keeping track of moves made by readers during case study. This data was a part of evaluation criteria for the display study as well as be valuable reference for understanding psychophysical attributes in different display designs. During a reading session, the computer periodically (at the rate of every 5^{th} millisecond) recorded reading status, including anatomic location of displayed slice, displayed slice volume in the case of stereo or MIP mode, indication of status between observing case and scoring nodule, and indication of compositing methods between averaging and MIP in the case of stereo mode.

Case Selection (task 7)

The images selected for developing and evaluating the display were from subjects who have previously been scanned as part of the Pittsburgh Lung Cancer Screening Study, an ongoing lung cancer screening trial in our facility, of subjects that are considered be at high risk due to age and smoking history. The cases were acquired from a helical CT scanner (LightSpeed Plus, GE Medical Systems, Milwaukee, WI) with slices reconstructed at a thickness of 2.5-mm, which resulted in the production of, on average, 100 slices per case. The images were post-processed with a kernel convolution algorithm using GE standard software, to adjust image sharpness to be suitable for viewing lung tissue.

The cases we selected were from a population where lung nodules, of which most are noncancerous, are prevalent due to subjects' age and smoking history. The average number of nodules per case is about ten. The actual distribution of nodules in our cases produces a lower sensitivity, for case-based analysis, than we were expecting. The high probability of finding positive cases would provide low sensitivity and less statistical power for comparing radiologists' performance of nodule detection across the different display modes, if we conduct case-based comparison study. To enhance the analytical power of the data, we adopted an observation-based strategy as the primary method of analysis. That was, the performances were analyzed based on each finding in a case and analyzed by FROC or ROC with finding-based instance methods. The change in emphasis from a case-based study to an observation-based study means a significant increase in statistical power.

Meanwhile, we have included more readers in both the pilot study and the main study. Particularly, we have more readers, especially recruiting fellows, so as to reflect a spectrum of various levels of clinical experience of readers. This is an important issue for us, as we need to know if any performance change has any association with subjects' experience and, to the less extent, subjects' physical conditions (such as age, visual conditions), typically when stereoscopic vision and eyewear are involved in the study. The number of readers is increased from 3 to 6 in the pilot study, and from 6 to 8 in the main study. This increase of the number of readers gives further justification of reducing the number of cases we initially proposed.

As stated above, redefining the definition of study instance and adjusting reader structure allow achieving at least the same analysis power but fewer cases. As the result of that, the final number of cases we collected for this project is reduced to 290, which contain total of 1630 consensus nodules, in which 18 are highly suspicious and 10 are pathologically proved cancer.

a. Nodule verification

Like other cancers, biopsy for lung cancer is reserved for highly suspicious ones, but unlike other cancer diagnosis, lung nodules are highly prevalent in the screening population and most of the nodules are often indolent. Thereby, most detected lung nodules are left for followup without immediate biopsy. In our collected cases, only 10 out of 1630 nodules were pathologically proved to be malignant and the rest of the nodules are diagnosed based on clinical impressions. To obtain truth of nodule characteristics, we used disease free follow-up for negative nodule verification and meantime we adopted the consensus method, which is widely used and accepted in the scientific groups for lung nodule verification [10].

To obtain consensus results, three experienced thoracic radiologists have reviewed all the nodules. The review process included nodule identification, verification, and characterization. Such review process was repeated at least once to ensure agreement on nodule identification and characterization. Table 3 lists all fields of a nodule description.

Table 3.

Name	Specifications
Nodule number	Number of Nodules
Nodule size	Measured in x, y axes (mm)

Nodule primary character	solid; non-solid (ground glass); mixed solid and non-solid
Nodule borders	smooth; spiculated; lobulated
Calcifications	absent; present
Location (side)	right lung; left lung
Location (lobe)	Right lung: RUL; RML; RLL Left lung: LUL; LLL
Risk level	Probability of malignancy

b. Cases anonymization by honest broker

To comply with HIPPA regulations, the case collection process was solely conducted by an honest broker, who has no any association with this project. After the cases had been selected, all personal information were removed from the cases and the cases were renamed with study numbers.

Reader Training (task 8)

Participating radiologists receive an "Instructions for Observers" form for review (Appendix C), and the definition of abnormalities are discussed with each reader prior to both the pilot and main studies. Readers were also trained on the use of our scoring mechanism during training session. To verify that observer performance is not affected by relative unfamiliarity or familiarity with our system, we did performance comparison of each radiologist at the beginning, middle, an end of the study once we complete the data collection from the pilot study.

Perform Pilot Study (task 9)

The pilot study was organized as a retrospective study of 108 nodules in 30 cases. Six radiologists, 4 experienced radiologists and 2 fellows, have participated the study. The reading data were collected and analyzed for the performance. The analyzed results provided an opportunity to further refine other aspects of the definitive study.

Analysis of Pilot Study (Task 10)

There were a total of 286 nodule-like features found in the pilot study. Since each study case was interpreted in 3 display modes by 8 radiologists, any nodule-like feature in the study case could be found 24 times if the feature were detected by all of the radiologists in all of the display modes. Figure 2 shows the distribution of number of times features were detected, for example the leftmost bin represents the features that were found only by one radiologist in one display mode, so these features were the least agreeable ones; whereas the rightmost bin represents the features were found by all of the radiologists in all of the display modes, so these features were the most agreeable ones. This distribution gives general idea of the variability of inter-readers and inter-modes for lung nodule detection.

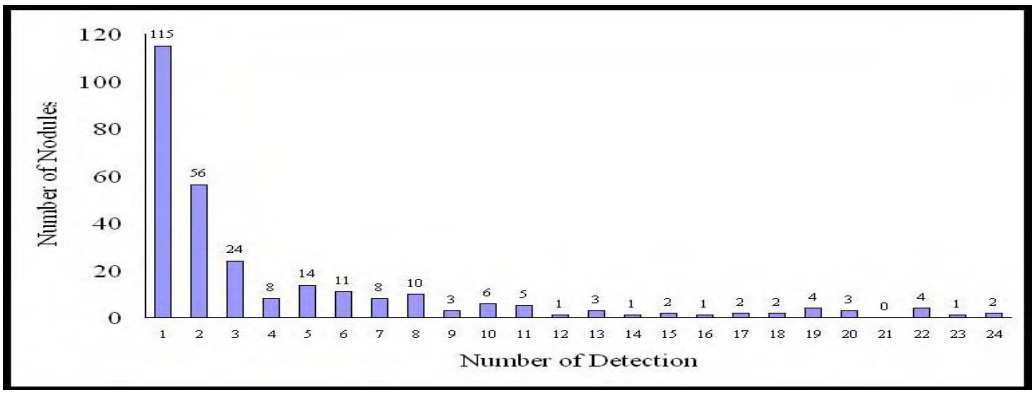


Figure 2. Distribution of number of times features were detected.

To reach a consensus result for nodule verification and nodule truth profile, we pooled features detected from the eight radiologists' interpretation in the three display modes. These features were reviewed and verified by an experienced chest radiologist, who did not participate the study but had read and discussed the cases with other radiologists multiple times.

We have also analyzed size distribution of the features found in the study, including the true positives and the false positives. Most of the features found in the pilot study are less than 10-mm as shown in figure 3.

Free-response Receiver Operating Characteristic (FROC) analysis suggests that the stereo display resulted the performance that was better than the orthogonal MIP display, but was equivalent to (or slightly better than) slice-based display, although no statistically significant difference was shown between the three display modes. The figure-of-merit (FOM) from the outputs of the JAFROC software (JAFROC, Chakraborty and Berbaum) for the three FROC curves were 0.57 (stereo display), 0.56 (slice-by-slice display) and 0.52 (orthogonal MIP display), respectively.

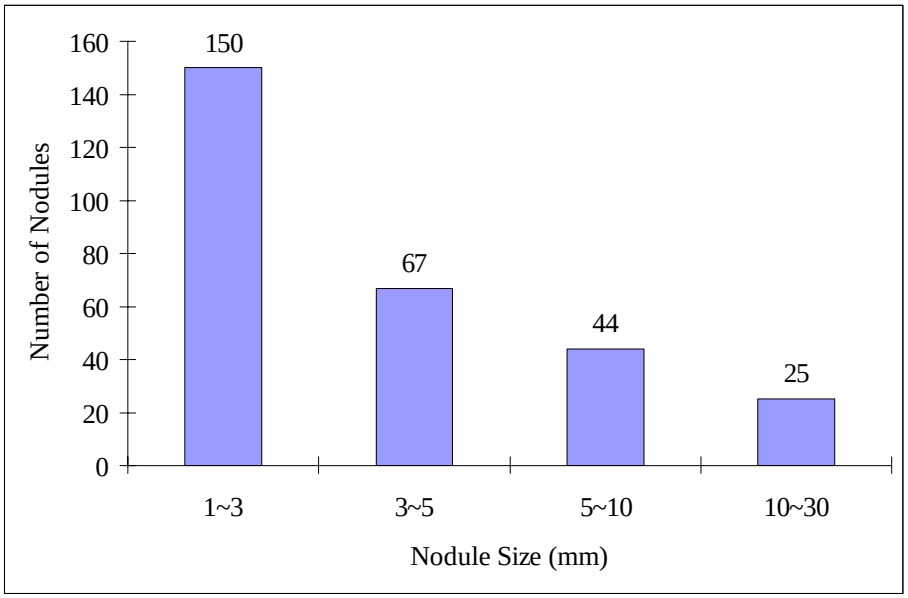


Figure 3. Distribution of nodule-like features by size.

One of the efficiency measurements is interpretation time on each tested display mode. We have recorded interpretation time as well as navigation patterns from 4 participating radiologists randomly selected to anonymize attributes associated with each individual. By averaging the time over the 4 radiologists on each display mode, we have shown that the average interpretation time was significantly less with the stereo display (3.5 minutes per case) than with the slice-by-slice display (4.5 minutes per case). There is, however, not much difference in interpretation time between in the stereo display and the orthogonal MIP display (3.7 minutes per case).

Even though the stereo display resulted generally less interpretation time and less falsely claimed nodules among the three tested display modes, the overall performance from the stereo display did not surpass the one from the slice-by-slice display. Subjective opinions and objective observations suggest that training effects significantly influence radiologists' search behavior and interpretation results. Of the three display modes tested in the pilot study, the stereo display has never been used or tried by the participating radiologists and the orthogonal MIP display has been experienced to a very limited extent. We observed from the navigation patterns recorded from 4 participating radiologists that, at the beginning of the study, radiologists were vigorously tuning their search patterns to try to find the optimal search pattern and optimal viewing volume with the two 3D displays. Towards the end of the study, the navigation patterns in either the orthogonal MIP display or the stereo display were getting much easier and smoother, and stayed in a more stable and controllable manner (see figure 1, 2 in Appendix B). In contrast, the navigation patterns in the slice-by-slice display were more randomized and undifferentiated between the beginning and the end of the study (see figure 3 in appendix II). Further, we observed that when interpreting cases with the 3D displays, radiologists tended to adjust the viewing volume from initial thick slab to single slice during early stage of the study. As single slice was the subset of the viewing volume in these volumetric display modes, the preference for the single slice suggested strong influence of training effect to radiologists' interpretation behavior.

Further analysis from the search patterns revealed that some of the missed nodules were actually received extra attention from the radiologists despite of no report being filed. There were about 25% of missed detections that received extra attention in the slice-based mode, 15% in the orthogonal MIP mode and 16% in the stereo mode. Radiologists have been trained conventionally in single projected radiographic image interpretation and maintained consistent practice manner for scrutinizing this kind of images carefully. But they are not extensively exposed to 3D display for volumetric data and, meantime, lacking of systematical training for volumetric data interpretation. Furthermore, since volumetric display can show more information in one view and clear geometrical relationship for easy understanding compared to single slice based display, radiologists may be over-confident for their observation and tend to neglect some subtle structures needed for more attention and/or different skills in 3D view. Appropriate training and practice, therefore, is necessary for achieving optimal performance with 3D display device and new display technology.

While novelty seemed to substantially affect navigation patterns and the performance, other factors associated with our 3D displays may also influence the results. Despite similarity in the navigation patterns and in the use of thickness information, the orthogonal MIP rendered display and the stereo display showed some differences in nodule detection. Vessel-like structures were much easier to be mistakenly recognized as nodules in the orthogonal MIP display as compared to that in the stereo display. Overall, the orthogonal MIP resulted more false positive findings than the stereo display (table 4) and the lowest performance score among the three display modes, although with no statistically significant difference. The low performance and high false positive rate of orthogonal MIP are most likely attributed to superimposed structures of monoscopic thick slab. Despite high contrast volumetric images, orthogonal MIP rendering may not produce correct geometric representation of volumetric objects due to that the algorithm takes the highest intensity along each projection ray, which may very well not preserve structural continuity between adjacent pixels in the rendered image. The stereoscopic rendering, on the other hand, was implemented with perspective transformation and transparency mechanism so that no superimposition was introduced and local geometric information was better preserved, especially with averaging method.

Table 4. Distribution of false positive findings in different structural groups.

	Stereo	**Orthogonal MIP**	**Slice by slice**	**Total**
Vessel	11	27	18	56 (32%)
Scar	23	32	33	88 (51%)
Other	10	7	12	29 (17%)
Total	44 (26%)	66 (38%)	63 (36%)	173 (100)

Subjective evaluation of the study program and data analysis indicate that the study design, including the number of readers, sample size, the scope of data collection and the study software, was satisfactory and did not need any major changes when applied to the main study. One minor adjustment for study software was to record time point when a nodule-like feature is detected and characterized, so that we can locate the detection activity during interpretation course and study search patterns for improving display design in the future.

Perform Main Study (Task 11)

We have conducted the main observer performance study that uses larger database, more readers, and improved study design based on the feedbacks from the pilot study. The main study was organized as a retrospective study of about 1154 nodules in 100 cases. Eight radiologists who have extensive experience in reading chest CT, have participated the study. Among the 8 radiologists, 4 of them were in the pilot study. Those who have not been in the pilot study had a short training course on the study procedures before the actual study. A total of 1154 suspicious lesions, including true and false findings, have been found during the readings from all readers in the 3 display modes. The initial findings from all readers and all display modes were first pooled together for consolidating and establishing consensus truth. Same lesions found in multiple display modes were identified by comparing and matching 3-dimensional locations.

Data analysis for main ROC study (Task 12)

The mean time for reading each case varied between readers for all cases and all display schemes, ranging from 0.88 ± 0.63 minutes to 5.50 ± 2.01 minutes. The mean time over all readers and examines were 2.78 ± 2.01, 2.71 ± 1.89, and 2.75 ± 1.92 minutes for stereo, MIP and slice-by-slice, respectively. There is no significant statistical difference of interpretation times between the three display schemes.

For stereo and MIP display schemes, time spent on reading in various slab thickness was recorded and expressed as percentage of the total reading time. Figure 3 shows the results of averaged times from the eight radiologists at each slab thickness point. The results indicate that most of the radiologists preferred reading 3D images with some volume but not in excessive thick volume, and there was no difference in the time distribution pattern between stereo and MIP schemes.

Averaged areas under ROC curves (A_z) for the performance on stereo, MIP and slice-by-slice displays are 0.67±0.06, 0.65±0.06, 0.65±0.04, respectively. Although there is a little better A_z for stereo than for MIP or slice-by-slice, there is no significantly statistic difference on performance between three display schemes.

The overall performance from the three display schemes was relatively low. One of the reasons for that is due to inter- and intra-reader variability. In this particular study, there were 8 radiologists and 3 display schemes; therefore, there should be 24 observations for each finding if all the radiologists found this one in all the 3 display schemes. When examining the results, we found that about 40% of the total findings were found by only one radiologist in one of the three display schemes (figure 4). Among those single-observation findings, about 37% of them are true positive findings and are equally distributed within the 3 display schemes, which implicate that for the 3 display schemes tested in the lung nodule detection, the diagnostic efficacy is equivocal. The equivocal efficacy was also indicated by A_z values of ROC curves and time measures.

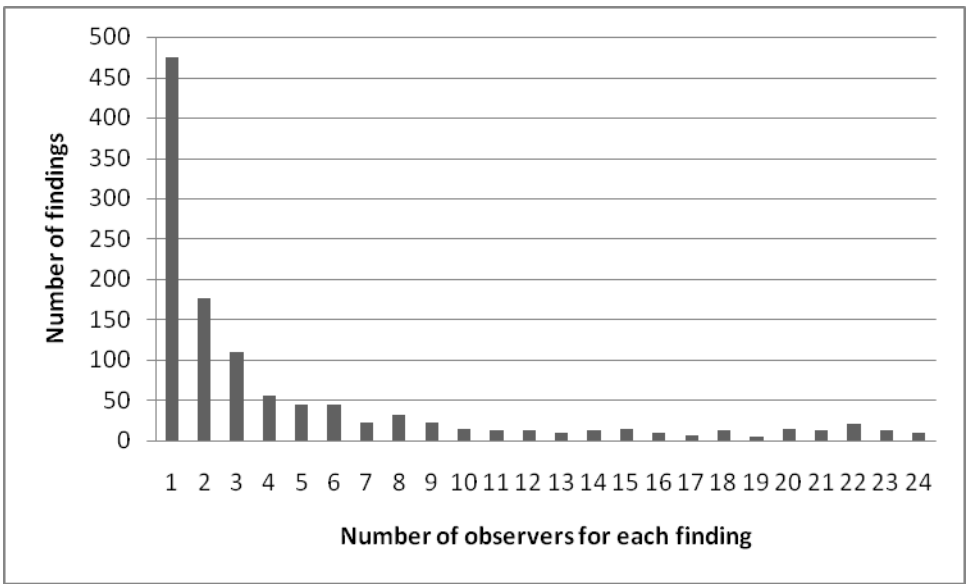


Figure 4. Distribution of all findings (positives and negatives) is graphed based on the number of observations. The maximum number of observations for each finding is 24 (the candidate nodule was observed by all 8 radiologists in all 3 display schemes) and minimum number of observation for each finding is 1 (the candidate nodule was found by only 1 radiologist in only 1 of the 3 display schemes).

At 100 slices per case used in this study, 3D displays seem not having much help to improve efficiency of reading and interpretation. At this level of CT resolution, it may not be easy to differentiate the level of efficiency of the 3 display schemes due to experience, training effect and information per slice. As showing in the times spent on slab thickness analysis (figure 5), radiologists still spent quite amount of time on thin slab or in one slice view despite of 3D displays. When the number of slice increased to over 500 due to the improvement of imaging techniques for higher resolution, which is currently the common case, it is hard to

imaging radiologist can continue practicing slice-by-slice interpretation. Volumetric displays can be more efficient in information retrieving, time and ergonomic issues, and can add value to efficiency.

A. B.

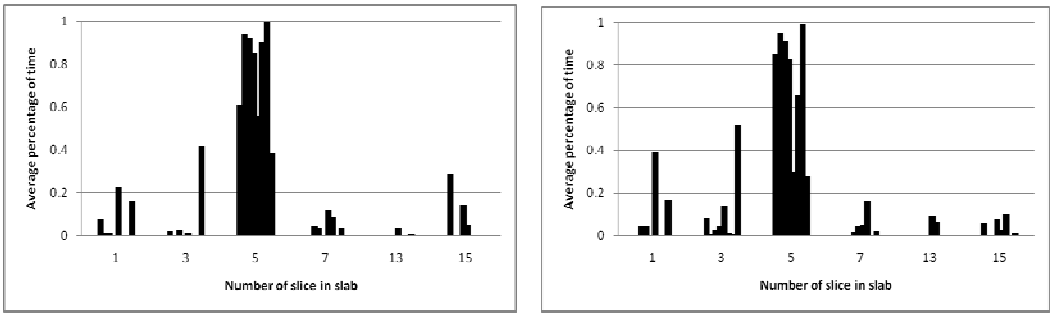


Figure 5. Average time, expressed as percent of total time, spent on different slab thickness for stereo display (A) and MIP display (B) schemes. At each slab thickness point, there are 8 plotted bars in which each represents time used by each of 8 radiologists

Analysis of supplementary issues (Task 13)

Issues, such as reader's use of slice thickness, window / level control, and subjective responses, are addressed and discussed in many of the above sections. Overall, radiologists were giving very positive feedbacks on 3D display schemes, especially the stereo display for its clear and non-overlapped 3D structures. The effective viewing thickness preferred by radiologists is very consistent between the pilot study and main study; that is most of the radiologists preferred reading 3D images with some volume but not in excessive thick volume. The results suggests that with increasing exposing, training and practicing in 3D imaging and 3D displays, radiologists can quickly move into new imaging technologies and effectively take advantage of3D image paradigms.

KEY RESEARCH ACCOMPLISHMENTS:

- Developed the mechanisms for a real-time stereographic display for volumetric datasets of lung CT images.
- Developed study software for investigating observer performance on stereo display and other commonly used displays for lung CT images.
- Developed 3D geometric projection algorithms and transparency-contrast models for compositing of stereo images from volumetric lung CT data.
- Collected 290 cases, which include 1630 nodules with 28 that are highly likely as malignancy.
- Trained readers for performing observer performance study.
- Performed the pilot study and collected the data; analyzed the data and gained some insight of interpretation behavior on volumetric displays that can be valuable for future guide of more efficient medical image rendering and display.
- Implemented rendering and display software on programmable graphics card that has achieved real-time volume rendering/displaying and user manipulation/interaction.
- Studied spectrophotometric characteristics of stereographic image pairs for better understanding stereo image quality and display issues.
- Completed a main study, and collected and organized the data from the study; the conclusion from the main study is stereo display produced better performance but lacking of statistic significance at the level of reading 2.5mm slice thickness and 100 slice /case; there are enough room for further improved in terms of rendering scheme, flexibility of volume thickness and cutting planes, and real-time image processing for better comprehension and easy maneuver of 3D images.

REPORTABLE OUTCOMES

Peer reviewed paper

Stereo CT image compositing methods for lung nodule detection and characterization, Academic Radiology 2005, 12:1512-1520.

Characterization of Radiologists' Search Strategies for Lung Nodule Detection: Slice-Based Versus Volumetric Displays, Journal of Digital Imaging, 2007; ISSN 0897-1889, page(s):1-11.

Real-Time Stereographic Rendering and Display of Medical Images With Programmable GPUs, Computerized Medical Imaging and Graphics, 2008; 32:118–123.

Display schemes for Lung nodule CT screening (manuscript preparation)

Presentation

Real-time stereographic display of volumetric datasets in radiology, 2006, SPIE, Electronic Imaging vol 6055, 1A-1 - 1A-6.

Stereo display of CT images for lung cancer screening: a pilot study, 2007, SPIE, Medical Imaging, vol. 6516.

Application of 3D Tensor Fusion for Segmenting Cylindrical Segments and Bifurcations from Volumetric Datasets, 2007, SPIE, Medical Imaging, vol. 6512.

Photometric Correction of Stereographic Image Pairs, 2008, ICISH'08, China, Page(s):108 - 110.

Grant application

Optimizing MDCT display for detection and diagnosis of pulmonary embolism, Summated to NIH, June 1, 2005.

Real-time stereo projection and display for 3D radiographic data, Submitted to NIH, June 17, 2005.

Real-time interactive display for virtual endoscopy, Submitted to NIH, October 19, 2005.

Immersive "Wall-of-Images" Display for Radiology - Preliminary Assessment, Submitted to NIH, September 2006.

Real-Time Interactive Stereo Display of Breast Tomosynthesis, Submitted to Susan G. Komen Breast Cancer Foundation, October 2006.

Immersive Stereographic Display for Real-Time Navigation through 3D Datasets, Submitted to NIH, October 2006.

Immersive Environment for High-Quality, Portable Display of 3-D Radiographic Datasets, Submitted to NIH, September 2007.

Investigations of ROC and FROC in the presence of verification bias and uncertain, submitted to NIH, October 2008

CONCLUSION

Our primary objective is to determine whether a stereoscopic display concept has potential for improving the efficiency and accuracy of chest CT interpretation for lung cancer screening. During the funding period, we have investigated several issues related to stereo display for medical 3D images, specifically CT images for lung cancer screening.

The preliminary data from pilot study and the data from main study all showed that the stereo display overall has resulted better and more efficient performance for lung nodule detection. Novelty and training effect, however, could possibly affect efficiency of using the 3D displays and other new technology related to medical imaging applications.

Our results strongly suggest that systematic training and practice is necessary for achieving optimal performance with stereo display device and other 3D display technology in medical image diagnosis. Optimal rendering methods are tasks specific and can be the key step in optimizing stereographic displays. Further improvement can be done in terms of rendering scheme, flexibility of volume thickness and cutting planes, and real-time image processing for better comprehension and easy maneuver of 3D images.

The reader interpretation patterns revealed in the study, and other possible improvements of display software design based on the observations from our study can be generally applied for improving performance in medical image interpretation.

REFERENCES

1. Greenlee RT, Murray T, Bolder S, et al. Cancer statistics, 2000. CA Cancer J Clin 2000; 50:7-33.
2. Ries LAC, Elsner MP, Kosary CL, et al. SEER: cancer statistics review, 1973-1997. Bethesda, MD: National Cancer Institute, 2000.
3. van Klaveren RJ, Habbema JDF, Pedersen JH, et al. Lung cancer screening by low-dose spiral computed tomography. Eur Respir J 2001; 18:857-866.
4. Flehinger BJ, Kimmel M, Melamed M. The effect of surgical treatment on survival from early lung cancer. Chest 1992; 101:1013-1018.
5. Sobue T, Suzuki T, Matsuda M, et al. Survival for clinical stage I lung cancer not surgically treated. Comparison between screen-detected and symptom-detected cases. Cancer 1992; 69:685-692.
6. Miettinen OS. Screening for lung cancer. Radiol Clin North Am 2000; 38:479-486.

APPENDICES

Appendix A
Peer reviewed paper: **Stereo CT image compositing methods for lung nodule detection and characterization**.

Appendix B
Peer reviewed paper: **Characterization of Radiologists' Search Strategies for Lung Nodule Detection: Slice-Based Versus Volumetric Displays**, Journal of Digital Imaging, 2007; ISSN 0897-1889, page(s):1-11.

Appendix C
Peer reviewed paper: **Real-Time Stereographic Rendering and Display of Medical Images with Programmable GPUs**

Appendix D
Conference paper: **Real-time stereographic display of volumetric datasets in radiology.**

AppendixE
Conference paper: **Stereo display of CT images for lung cancer screening: a pilot study.**

Appendix F
Conference paper: **Photometric Correction of Stereographic Image Pairs.**

SUPPORTING DATA
See APPENDICES, and the tables and the figures in the BODY section.

Appendix A

Stereo CT Image Compositing Methods for Lung Nodule Detection and Characterization

Xiao Hui Wang, Walter F. Good, Carl R. Fuhrman, Jules H. Sumkin, Cynthia A. Britton, Saraswathi K. Golla

Department of Radiology, University of Pittsburgh,

ABSTRACT

Rationale and Objectives Stereographic display has been proposed as a possible method of improving performance in reading CT exams acquired for lung cancer screening. Optimizing such displays is important, given the large volume of image data that must be evaluated for each of these exams. This study was designed to explore certain tradeoffs between rendering methods designed for the stereo display of CT images.

Materials and Methods Stereo CT image compositing methods, including distance-weighted averaging, distance-weighted Maximum Intensity Projection (MIP) and conventional MIP, were applied to lung CT images, and compared for lung nodule detection and characterization.

Results The Jonckheere test indicated that there was a statistically significant ($p<0.01$) increase in contrast among the three compositing methods. The Wilcoxon-Mann-Whitney test showed significant differences in contrast between distance-weighted averaging and conventional MIP ($p<0.01$) and between averaging and distance-weighted MIP ($p<0.05$), but not between distance-weighted MIP and conventional MIP ($p>0.05$). The conventional MIP compositing was found to provide the highest image contrast but produced ambiguities in local geometric detail and texture while the averaging resulted in the lowest contrast but preserved geometric detail. Distance-weighted MIP partially recovered geometric information, which was lost in the images composited with conventional MIP.

Conclusion Our results indicate that distance-weighted MIP may be a better choice for nodule detection in stereo lung CT images for its high local contrast and partial preservation of geometric information, while compositing by distance-weighted averaging is preferable for nodule characterization. The relative clinical value of these compositing methods needs to be further evaluated.

INTRODUCTION

Lung cancer is a leading cause of cancer death in both men and women in the United States [1], [2]. It has been shown that early detection and treatment can effectively reduce mortality from most types of lung cancers [3], [4], [5], [6]. In several large lung cancer screening trials, low-dose helical computer tomography (CT) has proven to be an effective tool for lung cancer screening with superior sensitivity (80 ~ 90%) for early detection, compared to other methods (e.g., chest radiographs with about 23% sensitivity and sputum cytology with 10 ~ 20% sensitivity) [3], [7], [8], [9], [10]. Despite the advantages of using CT as a screening tool for lung cancer, current methods for displaying and interpreting chest CT data are inefficient and inadequate.

At present, the two most common viewing methods employed by radiologists for interpreting these studies involve either reading individual images in a sequential slice-by-slice mode, or viewing thicker slabs comprised of multiple sequential slices projected onto a 2D display. The slice-by-slice method makes it necessary for radiologists to mentally reconstruct 3D information represented in sequential 2D slices in order to differentiate between nodules and linear structures, such as blood vessels, passing through the slices. In addition, the signal-to-noise ratio in single slices may be too low for subtle lesions to be reliably detected.

The use of thicker slabs, obtained by combining thin slices and then projecting this volume onto a 2D display, partially solves these problems but introduces other issues. The thicker the volume being projected, the more likely it is that the superimposed tissues in the projection will result in ambiguities in identifying structures within the projected volume. Furthermore, if slabs are projected by some form of averaging of all voxels projected onto a pixel, the contrast of small nodules may be reduced and this can result in a corresponding reduction in detection performance [11]. As combined slices become thicker by adding more of the originally acquired thin slices the contrast of smaller nodules, which often are visible on only one or two thin slices, may be reduced by averaging their voxel values with voxel values in slices not containing the nodules.

Over the last decade, maximum intensity projection (MIP) has gained extensive attention for its effectiveness in projecting thicker volumes of CT data, especially angiographic image data [12][13]. This method was originally introduced to overcome the contrast reduction associated with averaging methods and has been studied extensively for this purpose [11]. MIP orthogonally projects the volume of the combined slices onto a 2D display by using the maximum pixel intensity along each ray of the projection. In cases where the object being sought is brighter than superimposed tissues, then the object will be visible in the MIP image. While MIP can improve contrast in volumes that are not too thick, as the thickness of the volume increases the probability that some overlying tissue will have a higher voxel value increases, and this can reduce contrast in MIP images.

MIP is commonly used in procedures requiring the visualization of vascular structures because of its ability to clearly delineate bright objects overlaying a darker background, without significant loss of contrast. In general, MIP displays provide a mechanism for observers to specify the number of individual slices to be combined to form each thick slab, and then enable the observer to move this virtual slab through the data volume.

While such a rendering technique provides no information of depth or of the geometrical relationships between structures (i.e., tissue superposition), albeit the appearance of 3D projections, it does reduce the number of images that must be viewed in order to cover the entire 3D volume. However, the images it produces are, in principle, in conflict with plausible psychophysical transparency/brightness and geometric vision models. Note that, in order to make it possible for observers to appreciate the 3D structure of displayed volumes, these displays sometimes allow the volume to be continuously rotated, though it is known that this kind of image motion actually reduces the effective image resolution as perceived by viewers.

It is likely that higher resolution CT imaging, which will produce more and thinner slices, will be adopted for lung cancer screening in

the future, and this will exacerbate problems associated with large data volumes as well as those due to limited signal-to-noise ratios in single slices. If helical CT is to eventually be used for screening for lung nodules, as many have suggested [14], [15], then the efficiency and accuracy of the interpretation process must be increased.

We have proposed the use of stereoscopic 3D displays for reading lung CT images as a possible means to alleviate the aforementioned problems, in hope that this will increase both efficiency and detection performance beyond what can be achieved with other display methods [16], [17]. Such displays, which have the potential to provide more natural representations and volume-based views of structures for viewers having normal binocular vision, have been studied in the past for the display of medical images under various circumstances but, due to technical limitations (e.g., computational power, display technology), these methods have not been widely adopted. However, the current trend toward acquisition of large 3D datasets combined with improvements in the relevant technologies for stereographic display make it prudent to reconsider the potential role of these displays for radiology, at the present time.

Lung CT images are well suited for stereo viewing because, by making air transparent, the sparsely distributed lung tissues of interest (e.g., vessels, airways and nodules) can be easily visualized. Stereoscopic displays have the potential to increase efficiency and signal-to-noise ratios by enabling the display of thicker tissue volumes but they do not introduce the tissue superposition ambiguities that are often associated with a monoscopic presentation of thicker tissue volumes.

For CT data to be displayed stereoscopically, a 3D dataset must be projected onto two views, corresponding to observers' left and right eyes, using a geometric perspective transformation which projects rays through the 3D volume onto individual pixels. This process produces images that simulate what each eye would see if the viewer were looking directly at the 3D dataset. The algorithm that calculates a pixel value from intensity values along a ray, a process often referred to as compositing, determines contrast in the final projected image. Choosing a compositing method that is optimal for a particular task is the key decision that must be addressed in optimizing stereographic displays.

Compositing methods are generally subdivided into those that are related to volume rendering (i.e., methods capable of showing internal structure) and those based on surface rendering. Volume rendering is the most appropriate method for nodule detection in lung CT images, due to its ability to represent internal voxels and to the fact that no segmentation of surfaces, which could introduce artifacts, is required. Within the general paradigm of volume rendering, there are three compositing methods that appear to have some applicability to rendering chest CT data. These include distance-weighted averaging, conventional MIP and distance-weighted MIP.

For most applications, averaging voxel values along rays within the projected volume (i.e., averaging several CT slices to create a thicker image), with or without some form of distance-weighting, is the most commonly used method for compositing stereo volumetric images from individual slices, owing to its preservation of the internal properties of objects. The drawback of the averaging method, as discussed above, is the decrease in contrast of small objects with increasing volume.

Because of its ability to largely solve the contrast problem in many situations, MIP has been widely adopted for the monoscopic display of thicker slabs. However, it is not *a priori* clear that MIP is suitable for stereo compositing because of its inability to preserve the local geometric features and texture of objects and, to a large extent, it is this local information that is required for stereopsis. To achieve stereopsis from two views, a viewer needs to detect corresponding features in each view and to determine the relative geometric disparity between those features. Features that appear in one view (i.e., are a maximum along projections onto the view) may not be represented in a different view if they are not of maximum intensity along a projection for that view, and such features may provide misleading geometric disparity information. This causes small bright objects to have a specular appearance. However, MIP does preserve the presence, but not the exact geometry, of sparsely distributed clusters of bright voxels when they are viewed against a darker background, which is essentially the situation that usually occurs when a nodule is displayed in a thick slab from axial CT slices of the lung. For these reasons, images produced by MIP are, in principle, in conflict with plausible psychophysical transparency/brightness and geometric vision models.

Distance adjusted MIP is an attempt to combine conventional MIP with a psychophysically plausible depth cue, while retaining the contrast advantages of MIP. In this method, the brightness of voxels is reduced as a function of their depth within each projected volume. Brightness as a function of depth is a cue that most individuals subconsciously use all the time. In viewing objects in any transparent medium that has a constant optical density per unit of thickness, the brightness of an object will vary according to an exponential function of depth. MIP displays using this sort of depth cue vary the brightness throughout slabs as they are moved through the tissue volume to provide a plausible simulation of the optical properties corresponding to uniform optical density. This process has the added benefit for MIP that it can change the relative brightnesses of voxels based on the position of slabs containing the voxels, which can improve the detectability of objects that are similar in brightness to their backgrounds.

These various alternatives for compositing have not been previously compared in the context of nodule detection in chest CT, and the degree to which the concerns mentioned above could affect tasks such as nodule detection in lung CT images is unknown. This paper explores these compositing methods with the goal of identifying methods suitable for the stereographic display of lung CT for the purpose of nodule detection.

MATERIALS AND METHODS

1. Lung CT images and nodule verification

Ten sets of lung CT images, each containing a consensus-proven solitary lung nodule, were used in this study. For each image set, stereo image pairs were projected using four compositing methods including distance-weighted averaging, MIP and two versions of distance-weighted MIP. The images were displayed with window/level setting of 1800/-600 for MIP images and 1500/-600 for averaging. Contrast and nodule characteristics were compared between these compositing methods.

1.1. Lung CT images

Helical CT images were obtained from a low-dose CT lung cancer screening project conducted in the Medical Center, University of Pittsburgh. The CT cases were performed on a LightSpeed Plus multislice CT scanner (GE medical Systems, Milwaukee, WI) using X-ray tube current of 40-mA, voltage of 140-kVp and 0.5-mm pitch. The images were acquired in the axial plane and reconstructed to a thickness of 2.5 mm/slice with GE standard convolution software for lung tissue. The pixel size in each slice is 0.75-mm × 0.75-mm.

1.2. Nodule verification

All the nodules used in this paper were identified, verified, marked and characterized by three experienced radiologists. The verification process was repeated at least one time to ensure agreement on the

nodule identification and characterization. Table 1 lists the nodules and their properties.

2. Stereo compositing methods

The voxel shape in the CT images used in this study was nonisotropic in that images had been reconstructed to a larger thickness in z direction (i.e., slice thickness of 2.5-mm) than in x and y directions (pixel dimension: 0.75 × 0.75-mm). To approximate isotropic voxels, three slices were created by trilinear interpolation between each pair of adjacent CT slices. Consecutive slices, including CT and interpolated slices, within a given volume were used for generating a stereo pair.

In the conventional geometric perspective transformation that was used to compose left-eye and right-eye image pairs, we adopted an interpupilary distance of 6.5-cm, a viewing distance (the distance between eyes and the screen) of 45-cm and a display area of 25-cm × 25-cm. The perspective transformation was symmetrical, based on the assumption that a viewer is centered in front of the screen.

Each set of slices was composited to a stereo pair for each of the projection methods. Since the effect of superimposed tissue is different for each of the compositing methods, the characteristics of a projection method depend on the depth of a nodule within a 3D volume. Thus, in order for us to compare the depth-dependence of compositing methods, for each nodule we constructed three volumes with the nodule at different depths in each volume. Specifically, the three volumes were comprised of: 1) precisely those slices that cover the nodule; 2) all slices in the first set plus an additional 3 slices in front of the nodule; and, 3) all slices in the first set plus an additional three slices behind the nodule.

2.1. Traditional Compositing by Distance-Weighted Averaging

For this compositing method we adopted a light emission/transmission/occlusion model that assumed that each voxel emits light in proportion to its brightness when the CT images are displayed at a normal window and level for nodule detection, and that uses distance information (distance weighing factors) to determine the amount of this emitted light that reaches the projection plane.

Specifically, it was assumed that each slice has a fixed optical density that reduces the brightness of slices lying behind it. The total of all distance-weights was equal to one. The ratio of the weights between the last slice (the slice with the largest distance from screen) and the first slice (the slice at screen level) controls level of transparency for a given volume.

We have studied a range of these ratios for lung CT images, and empirically set the ratio to 0.5 in order to achieve a balance between the use of brightness weighting as a depth cue and the visibility of the back slice. The final value for a voxel is the sum of distance-weighted pixel values in a perspective transformation ray. The detailed calculations were described in references [16] and [17].

2.2. Stereographic MIP Compositing

This compositing method uses a perspective projection in which the maximum value along each ray is used as the projected value. Because the projected voxels may be different between the stereo views, it is not generally possible for an observer's vision system to unambiguously match corresponding points between the two views. This may cause small objects, such as lung nodules to have a speckled appearance. Thus, this method is not *a priori* suitable for stereo projection. In practice, we have found that the ambiguity in matching corresponding points between views primarily affects fine detail and does not interfere with the detection of objects that are comprised of large clusters of bright voxels, though the exact shapes of these objects cannot be unambiguously determined.

2.3. Distance-Weighted Stereographic MIP

In an attempt to incorporate a geometric cue common in traditional stereo projection methods, but also preserve the contrast advantage of MIP, we employed a perspective projection in which the maximum along each ray is weighted based on distance. Since MIP only takes one pixel value along a ray, the sum of total weight is irrelevant in this case and therefore, only the maximum (nearest) and the minimum (farthest) weights were empirically determined. Based on transparency/occlusion model, the first slice was weighted as 1, the last slice was weighted as 0.5 and weights for slices between the first and the last were calculated using a geometric sequence, assuming a fixed optical density for each slice.

There are two paradigms for applying distance-weighting to MIP. Voxel values can either be adjusted by distance weights prior to acquiring the maximum value along a ray (Distance-MIP), or alternatively, the maximum value can be chosen first and then adjusted by applying the distance-weighting factor (MIP-Distance). In this study, we have tried both strategies and a comparison was made based on image appearances and contrast measurements.

3. Contrast analysis

The boundaries of the identified nodules were manually marked on CT slices. The background included the area of 20-mm from the nodule boundary in x, y and z directions. Local contrast was measured using Michelson Contrast measure C_m, with local maximum intensity L_{max} and minimum intensity L_{min},

$$C_m = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}}.$$

The one-tailed Jonckheere test was used for testing for a significant trend of increasing contrast for the group of three compositing methods as ordered in the statement of the hypothesis. The one-tailed Wilcoxon-Mann-Whitney Test was used for testing the significance of difference in contrast measures between averaging and conventional MIP, between averaging and distance-weighted MIP, and between distance-weighted MIP and conventional MIP.

RESULTS

We have compared averaging, MIP and distance-weighted MIP applied to various nodule types, sizes and locations. Figure 1 illustrates three sets of nodules to show typical examples resulting from the three stereo compositing methods. Figure 1A shows a solid nodule with smooth border, and Figures 1B and 1C show nonsolid nodules with spiculated borders. As is shown in Figure 1, MIP and distance-weighted MIP produced higher local contrast than compositing by averaging, in all subsets. If the area surrounding a nodule contains dense structures, such as bone in this case, a nodule can be camouflaged by, or blended into, the background when stereo pairs are rendered with conventional MIP, as shown in Figure 1B and Figure 1C. This camouflage effect is less noticed in the distance-weighted MIP images.

Local contrast measures for 10 nodules with different compositing methods are listed in the Table 2, where each cell has 3 sets of Michelson Contrast numbers. Two numbers in a set are for the left and right images in a stereo pair, and the first set of numbers are from the subset of images containing a nodule, the second set of the numbers are from the images containing the nodule plus extra front slices and the third set of the numbers are from the images containing a nodule plus extra back slices. In general, conventional MIP and distance-weighted MIP (both MIP-Distance and Distance-MIP)

produced higher contrast measures compared to the averaging method. The contrast measures from MIP method are fairly constant in all slab thickness used in this study, while the averaging method resulted in decreasing contrast as slab thickness increased. There is a statistical significant ($p<0.01$) of ordered contrast change (averaging, distance-weighted MIP, conventional MIP) among the three compositing methods as measured by Jonckheere test. Also notice that the MIP-Distance approach, in general, had slightly higher average contrast measures than the Distance-MIP approach. When comparing contrast between the compositing methods, the one-tail Wilcoxon-Mann-Whitney Test indicated that there was a significant difference between averaging and distance-weighted MIP (either MIP-Distance approach or Distance-MIP approach, $p<0.05$), between averaging and conventional MIP ($p<0.01$) in all three compositing thicknesses. There was no statistically significant difference in contrast measure between distance-weighted MIP and conventional MIP in all three compositing thicknesses ($p>0.05$), except one pair, which was the difference between Distance-MIP and conventional MIP applied to the subgroup of nodule with three front slices ($p=0.0375$).

Finally, we have compared the three compositing methods visually for nodule characterization. As shown in the Figures 1B and 1C, the spiculated nodule border is clearly visible in all stereo pairs composited with averaging method, while this characteristic is not preserved in some of the stereo pairs composited with conventional MIP method, especially those composited from thicker slabs. Distance-weighted MIP partially overcomes the problem of conventional MIP, and spiculated borders are still visible in the thicker slabs. As for the smooth border we observed similar phenomenon. The geometric relationship is well presented with gradient changes in intensities along the smooth border of the nodule (Figure 1A) in the stereo pair composited with averaging method. The same nodule is displayed with lack of geometric fidelity, as shown by the sharp edge of the smooth border in the stereo pair composited with conventional MIP.

DISCUSSION

Previously, we have shown that stereoscopic display for lung CT images can dramatically improve display efficiency and lung nodule visibility, relative to conventional slice-based displays. In this present study, we have further investigated the effects of several commonly used compositing methods on nodule representation in stereo CT images.

The results from this study, as well as from others, have shown that unlike averaging methods, which sacrifice contrast in order to take account of each voxel in a volume, the conventional MIP method is able to retain contrast in cases where the object being viewed includes voxels that are brighter than the superimposed tissue. As applied to the task of lung nodule detection, despite a lack of geometric fidelity, the conventional MIP images generally produce high local contrast that separates a nodule from its background, and therefore enhances detection performance.

The improvements in nodule visibility with the MIP method, however, do not apply in certain cases in lung CT images. For a nodule to be detected with conventional MIP, it must contain some voxels that are brighter than its background. The case of a nodule overlying a rib is a particular concern though it occurs relatively infrequently in projections of axial slabs, unless the nodule is very close to a rib or the slab being viewed is relatively thick. Most voxels in such a nodule will not be as bright as voxels in the rib and the nodule may be almost invisible. Examples from this study (MIP images in Figures 1B and 1C), in which the nodules are almost indiscernible when they are close to bone tissue, have demonstrated this effect.

This same potential problem occurs in monoscopic MIP displays that are already in wide use. The situation is actually somewhat improved in the case of stereo viewing. Traditional MIP employs an orthogonal projection, the rays of which do not change relative to the 3D volume as a slab is moved in the axial direction. This means that when a voxel of a rib and a voxel of a nodule fall on the same ray of a projection, then the two voxels will always be superimposed when both are contained within a slab. The stereographic projections we are using involve two perspective transformations for each axial position of a slab. The angles of rays passing through a nodule, in each of these projections are different so that if two voxels are superimposed along one ray they will not be along the other. Furthermore, as a slab is moved in the axial direction, the orientation of a projection ray passing through a particular voxel changes continuously, so that the sets of superimposed voxels change continuously. Thus, if a slab is moved slightly, an obscured nodule is more likely to become visible with stereographic projection as opposed to monoscopic MIP.

One benefit of the distance weighted MIP projection is that it can reduce the brightness of the background. As the position of a slab is changed, the relative brightness weighting factors between voxels at different axial positions will change and, in many instances, there will be an axial position where a nodule will appear brighter that a rib and this can increase the likelihood that an obscured nodule becomes visible.

Segmenting the ribs and spine can theoretically reduce this problem, but that process may generate its own artifacts. Segmentation would involve some sort of surface detection, and if the surfaces were perfectly smooth, that should not present any difficulties. However, any realistic surface detection algorithm must differentiate between roughness of the surface and nodules lying near the surface. If a computer could do this reliably then there would be no need for the radiologist to view the images. Consequently, we are particularly concerned about the risks of segmentation in precisely those cases where it could be of potential value.

As described in the methods, we have tried two different ways for distance-weighted MIP compositing, the Distance-MIP approach and the MIP-Distance approach. The two approaches produced slightly different results due to the fact that different maximum intensities were chosen. For distance-MIP, the maximum value I_{max} is dependent on weight distribution in j elements I_j,

$$I_{\max} = \max(I_1 w_1, I_2 w_2, I_3 w_3, ..., I_j w_j).$$

If weight distribution, $w_1, w_2, w_3, \ldots, w_j$, drops fast, weighted values away from the front are greatly reduced, and it is more likely that values nearer the front will be chosen. In this case, the final image may represent only a few slices in the front, and volume information could be lost. In this study, that effect was not apparent because of the relatively slow change in the weighing factors. For the MIP-Distance approach, although it may not necessarily have the maximum intensity in the final images as might be the case if a chosen voxel had a large reduction in intensity because it was located in the back, it does take into account the entire volume and produces contrast generally as good as that with Distance-MIP. An example is shown in Figure 1B (3 slices + nodule) in which the geometric relationship was better preserved in the images rendered with MIP-Distance compositing compared to Distance-MIP composting. For the purpose of detection and diagnosis, the MIP-Distance approach may be preferred to the Distance-MIP approach.

Our results indicate that both versions of distance-weighted MIP partially recover geometric information lost in conventional MIP, by incorporating a distance cue into the compositing. As shown in the Figure 1, distance-weighted MIP reveals the nodule within the bone

area (Figures 1B and 1C), where the same nodule is not easy to detect in the conventional MIP projection. However, the distance-weighted MIP had image contrast nearly equivalent to that produced by the conventional MIP method. Overall, distance-weighted MIP retained most of the contrast of MIP, while improving geometric fidelity.

In addition to detection, the nodule characteristics are essential and critical for clinically differentiating benign from malignant. Although the conventional MIP is superior for detection, it was outperformed by the distance-weighted averaging method in terms of characterization, as can be seen in our results. The lack of fidelity in nodule shape and geometric representation, in the conventional MIP images, are attributed to the nature of MIP compositing, in which the two views may be based on projections of different voxels. The averaging method, on the other hand, has been shown in this study to faithfully retain the characteristics of the nodules, including structural, spatial and geometric information. This can be observed in all nodules tested in this study.

Because nodules must be detected before they can be characterized, and we have shown a tradeoff between contrast for detection and geometric fidelity for characterization, we believe that the use of two separate stereoscopic display modes is both desirable and feasible. First, the image data should be displayed using some form of distance-weighted MIP to maximize detection performance. Once a nodule has been detected, the thickness of the displayed volume can be adjusted so as to include only those slices that contain the nodule, and this volume can be displayed using an averaging compositing method in order that the nodule can be more accurately characterized.

Unfortunately, while monoscopic approaches to 3D display can be appreciated by most individuals, there is a significant variation across the population in the ability to achieve stereopsis [18]. It has been reported that 2% to 4% of individuals are stereopsis blind and another 10% to 15% have a stereopsis deficiency in the sense that they have difficulty deriving 3D information from random-dot stereograms. Much of the stereopsis impairment, reflected in these numbers, has been attributed to strong uncorrected astigmatism [18]. This would seem to limit the value of stereo display to the subpopulation having normal binocular vision. This kind of limitation is not unprecedented in radiology – colorblind radiologists would likely have difficulty taking advantage of the benefits of using color on a PET display, for example, and previously, medical student applicants for a stereo X-ray fluoroscopy training program have been prescreened using random-dot stereograms. Nevertheless, for the majority of radiologists, stereographic methods may provide a more natural way for them to perceive the spatial relationships in 3D volumetric datasets, and those radiologists who are unable to achieve stereopsis would be able to continue to reading from traditional monoscopic displays or employ other means of representing 3D data.

There is also a concern that even individuals who can achieve stereopsis on surface rendered images will have difficulty seeing volume rendered images in stereo, though there is no direct evidence of this. In any event, in the scenario we are studying with respect to stereographic projection of the lung, the objects whose positions we are trying to clarify are small relative to the lung volume and tend to be sparsely distributed. In the detection task, it is not as important that we see the internal structures of the objects in stereo, as it is that we see the relative positions of the objects in stereo. We could achieve our goals by surface rendering the interior of the lung, but as mentioned above, that kind of calculation requires that rather intelligent emulation of the mental processing typical of radiologists, be performed in software. This has the risk of introducing other artifacts, and is not necessary to solve the detection problem.

The advantage of applying stereoscopic technique over other 3D rendering techniques for medical 3D data display is that stereo display employs a mechanism naturally used by human visual system for detecting and characterizing objects. By presenting data in a volume-based stereoscopic format, radiologists' efficiency and accuracy in interpreting CT images may be significantly improved. Although the actual value of stereo display for lung cancer screening is not yet known, it was the intention of this study to begin to investigate methods for achieving optimal image quality in anticipation of future observer performance studies aimed at measuring the efficacy of stereo displays for chest CT. It is noted that this study has involved a small sample size of test sets, and a much larger study would be required to clarify all differences between the compositing methods.

ACKNOWLEDGEMENTS

This work is sponsored in part by grant CA80836 from the National Cancer Institute, National Institutes of Health, and also by the US Army Medical Research Acquisition Center, 820 Chandler Street, Fort Detrick, MD 21702-5014 under Contract DAMD17-02-1-0549 and contract PR043488. The content of the contained information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

References

1. Greenlee RT, Murray T, Bolden S, et al. Cancer statistics, 2000. CA Cancer J Clin 2000; 50:7-33.
2. Ries LAC, Eisner MP, Kosary CL, et al. SEER: Cancer statistics review, 1973-1997. NCI 2000.
3. van Klaveren1 RJ, Habbema JDF, Pedersen JH, et al. Lung cancer screening by low-dose spiral computed tomography. Eur Respir J 2001; 18:857-866.
4. Flehinger BJ, Kimmel M, Melamed M. The effect of surgical treatment on survival from early lung cancer. Chest 1992; 101:1013-1018.
5. Sobue. T, Suzuki T, Matsuda M, et al. Survival for clinical stage I lung cancer not surgically treated. Comparison between screen-detected and symptom-detected cases. Cancer 1992; 69:685-692.
6. Miettinen. OS. Screening for lung cancer. Radiol Clin North Am 2000; 38:479-486.
7. The American Society of Cytopathology, Cytopathological Practice Committee, "Nongynecological cytology practice guidelines." The American Society of Cytopathology, http://www.cytopathology.org/guidelines/nongynecologicalxi.php.
8. Jett JR, Midthun DE. Screening for lung cancer: current status and future directions. Chest 2004; 125(5 Suppl):158S-162S.
9. Kawahara M. Screening for lung cancer. Curr Opin Oncol 2004; 16:141-145.
10. Diederich S, Thomas M, Semik M, et al. Screening for early lung cancer with low-dose spiral computed tomography: results of annual follow-up examinations in asymptomatic smokers Eur Radiol 2004; 14:691-702.
11. Brown DG, Riederer SJ. Contrast-to-noise ratios in maximum intensity projection images. Magn Reson Med 1992; 23:130-137.
12. Napel S, Rubin GD, Jeffrey RB, Jr. STS-MIP: a new reconstruction technique for CT of the chest. J Comput Assist Tomogr 1993; 17:832-838.
13. Calhoun PS, Kuszyk BS, Heath DG, Carley JC, Fishman EK. Three-dimensional volume rendering of spiral CT data: theory and method. *Radiographics* 1999; 19:745-764.
14. Gohagan J, Marcus P, Fagerstrom R, et al. Baseline findings of a randomized feasibility trial of lung cancer screening with spiral CT scan vs chest radiograph: the Lung Screening Study of the National Cancer Institute. Chest 2004; 126:114-121.

15. Armstrong P, Husband JE, Holemans JA. Population screening for lung cancer. Hosp Med 2004; 65:404-411.
16. Wang XH, Good WF, Fuhrman CR, et al. Stereo Display for Chest CT. Proc SPIE 2004; 5291:17-24.
17. Wang XH, Good WF, Fuhrman CR, et al. Projection Models for Stereo Display of Chest CT. Proc SPIE 2004; 5367:676-686.
18. Julesz B. Foundations of Cyclopean Perception. The University of Chicago Press, 1971; page 270.

Table 1. Nodule information.

Nodule #	Size (mm^2)	Border	Characteristic
1	4×2	Smooth	Solid
2	8×6	Spiculated	Non-solid
3	10×7	Spiculated	Solid
4	7×7	Spiculated	Non-solid
5	6×4	Speculated	Mixture of solid and non-solid
6	6×5	Spiculated	Solid
7	5×4	Smooth	Mixture of solid and non-solid
8	10×9	Spiculated	Mixture of solid and non-solid
9	6×5	Smooth	Mixture of solid and non-solid
10	6×5	Smooth	Mixture of solid and non-solid

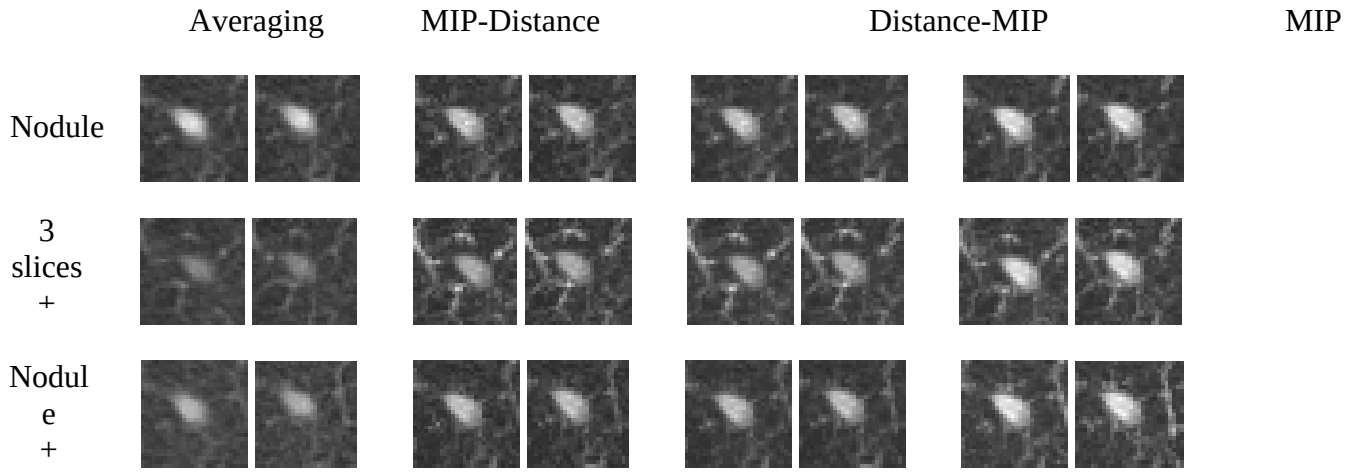
Table 2. Contrast measurements with Michelson Contrast Measure. Three sets of measurements in each cell are the stereo pairs (left and right) composited from the nodule slices, the nodule slices with 3 front slices, and nodule slices with 3 back slices, respectively.

Nodule #	Averaging	MIP-Distance	Distance-MIP	MIP
1	0.82 / 0.84 0.73 / 0.71 0.76 / 0.74	0.85 / 0.85 0.81 / 0.79 0.81 / 0.81	0.84 / 0.85 0.81 / 0.77 0.80 / 0.80	0.85 / 0.86 0.84 / 0.83 0.81 / 0.83
2	0.84 / 0.84 0.79 / 0.83 0.79 / 0.78	0.88 / 0.86 0.82 / 0.87 0.84 / 0.85	0.84 / 0.84 0.79 / 0.83 0.87 / 0.85	0.89 / 0.88 0.85 / 0.90 0.87 / 0.85
3	0.89 / 0.84 0.82 / 0.80 0.87 / 0.80	0.90 / 0.86 0.89 / 0.89 0.87 / 0.85	0.91 / 0.85 0.90 / 0.88 0.87 / 0.80	0.94 / 0.89 0.90 / 0.89 0.90 / 0.88
4	0.91 / 0.92 0.87 / 0.87 0.89 / 0.88	0.92 / 0.93 0.89 / 0.88 0.91 / 0.92	0.92 / 0.93 0.90 / 0.88 0.90 / 0.91	0.94 / 0.95 0.90 / 0.91 0.93 / 0.94
5	0.82 / 0.84 0.77 / 0.80 0.79 / 0.83	0.85 / 0.87 0.82 / 0.84 0.83 / 0.87	0.86 / 0.89 0.81 / 0.84 0.82 / 0.87	0.88 / 0.91 0.85 / 0.86 0.85 / 0.88
6	0.91 / 0.93 0.90 / 0.89 0.89 / 0.88	0.97 / 0.96 0.93 / 0.92 0.92 / 0.92	0.96 / 0.96 0.93 / 0.92 0.93 / 0.93	0.99 / 0.98 0.98 / 0.96 0.98 / 0.98
7	0.76 / 0.91 0.75 / 0.79 0.73 / 0.77	0.82 / 0.91 0.79 / 0.80 0.76 / 0.79	0.78 / 0.90 0.79 / 0.79 0.76 / 0.77	0.87 / 0.96 0.81 / 0.85 0.77 / 0.83
8	0.93 / 0.93 0.90 / 0.90 0.91 / 0.92	0.96 / 0.96 0.94 / 0.95 0.92 / 0.91	0.96 / 0.95 0.94 / 0.94 0.94 / 0.94	0.99 / 0.99 0.98 / 0.99 0.98 / 0.97
9	0.78 / 0.80 0.72 / 0.73 0.75 / 0.75	0.81 / 0.86 0.80 / 0.80 0.82 / 0.84	0.80 / 0.82 0.79 / 0.80 0.86 / 0.86	0.87 / 0.88 0.82 / 0.82 0.86 / 0.86

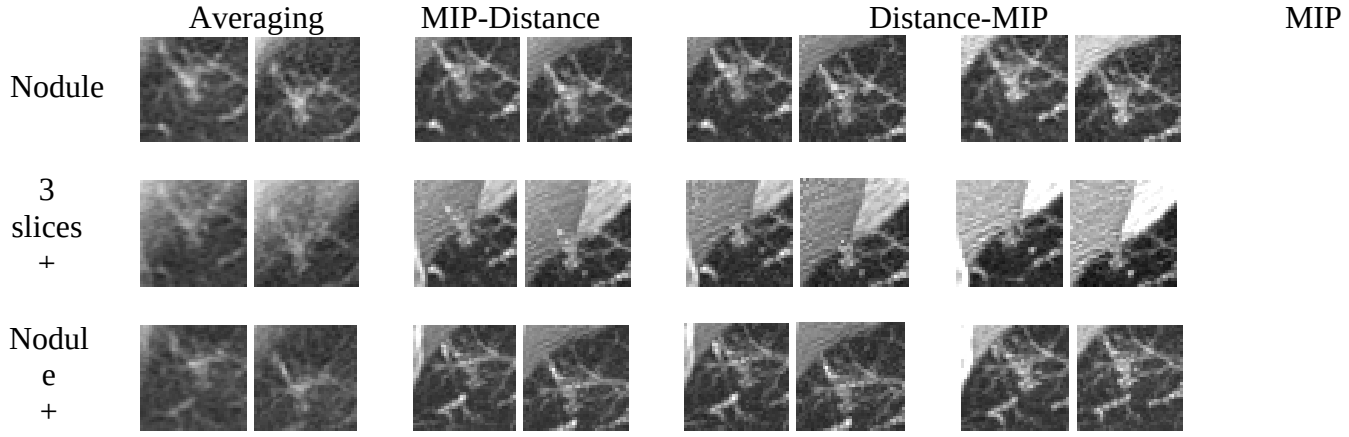
10	0.92 / 0.93 0.83 / 0.82 0.85 / 0.84	0.93 / 0.94 0.86 / 0.91 0.91 / 0.91	0.93 / 0.94 0.86 / 0.84 0.91 / 0.90	0.96 / 0.96 0.90 / 0.91 0.92 / 0.91

Figure 1.

A.



B.



C.

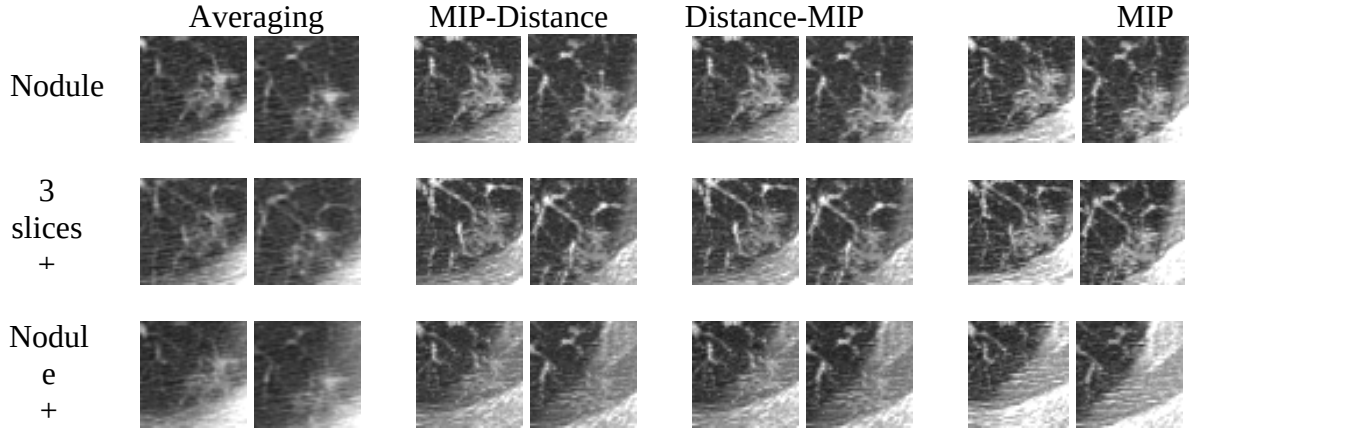


Figure 1. Stereo image pairs of three nodule sets (A., B., and C.). Images of A, B, and C are from nodule #1, nodule #2 and nodule #4, respectively as denoted in the table 1 and table 2. In the row labels of each set, "Nodule" means the images were composited from nodule slices, "3 slices + Nodule" means the images were composited from nodule slices plus 3 front slices, and "Nodule + 2 slices" means the images were composited from nodule slices plus 3 back slices.

Appendix B

Characteristics of CT Image Radiologists' Search Strategies for Lung Nodule Detection: Slice-Based Versus Volumetric Displays

Xiao Hui Wang, Walter F. Good, Janet E. Durick, Amy Lu, David L. Herbert, Saraswathi K. Golla, Kristin Foley, C, Samia Piracha, Dilip D.Shinde, Shindel , Carl R. Fuhrman, Cynthia A. Britton, Sherry S. Shang, Christoper Deible and Joan M. Lacomis

Department of Radiology, University of Pittsburgh, School of Medicine, Pittsburgh, PA

Abstract

OBJECTIVES. The purpose of this study was to investigate characteristics of radiologists' search patterns and search results in lung nodule detection on CT images with different rendering and display schemes for improving medical volumetric image visualization and diagnostic performance.
MATERIALS AND METHODS. Retrospective lung nodule detection with computerized tomographic images was conducted on three display modes, including slice-by-slice display, orthogonal maximum intensity projection display and stereoscopic display. Thirty lung-cancer-screening CT cases containing 91 nodules were used in the study, and eight radiologists interpreted the cases. Radiologists' search course within the volumetric data was recorded along with the probability of a nodule, location, size and shape for each detected feature. Characteristics of detected features and radiologists' search patterns were compared for the three display modes. The nodule detection performance was analyzed with Free-response Receiver Operating Characteristic method.
RESULTS. The stereo display provided better visual effect of 3D representation and produced better detection and classification performance with less interpretation time compared with other display modes tested in the study. However, the difference between the stereo display and the other displays was not statistically significant. Further analysis of the navigation patters showed that novelty and training effect were associated with the nodule detection performed on the volumetric displays. Among the tree display modes, the orthogonal maximum intensity projection display resulted the highest number of false positives, in which most were vessel structures. Scar tissue was the most common structure falsely recognized as lung nodule in all three display modes.
CONCLUSION. Our preliminary results indicate a potential role of stereo display for improving radiologists' performance in medical detection and diagnosis, and also strongly suggest that systematic training and practice is necessary for achieving optimal performance with volumetric displays or any new display technology in medical image diagnosis.

Keywords: volumetric dataset, navigation, stereoscopic display, lung nodule screening

Introduction

Medical image interpretation involves heavily human-image interaction. Extensive studies have been conducted to investigate eye search patterns on projected radiographic images for lesions [1, 2, 3, 4, 5, 6, 7]. The results indicate that the eye search characteristics are more on experience bases, are influenced by image quality, and can be correlated to the performance of detection and diagnosis [1, 3, 4, 6, 7, 8, 9]. These studies have helped to improve image quality, image representation and visual inspection technique. However, most of these works have been focused on 2-dimensional (2D) radiographic images and very little research by far has been done on human-computer interaction and searching behavior associated with 3-dimentional (3D) medical datasets.

Medical imaging is rapidly evolving into 3D representation [10, 11, 12]. In the near future, it is very likely that 3D datasets from various imaging modalities will dominate medical imaging format for diagnosis, treatment and image-guided surgery. Radiologists will have to adopt new search or navigation strategy to interpret image datasets. One big difference between 2D projected image and 3D image dataset is that resolution on each single 2D image in a 3D dataset is much lower than that on a single 2D projected image. Because of reduced resolution on each image and expanded information into one more dimension, radiologist needs to rely more on the information between images, which introduces information exploring in additional dimension and changes drastically the behavior of gazing and searching during image interpretation.

The features unique to 3D datasets are all likely to affect radiologists' interpreting behavior. For example, ever-increasing image volume forces radiologists to adopt computerized display (soft copy display) and to be involved more and more actively in computer-based procedures and operations for image interpretation. Furthermore, in order to optimally utilize information captured in 3D datasets, various computer algorithms have been developed to render 3D images into 2D images for display. Such practice has changed traditional radiographic presentation that radiologists have learnt and accustomed, and may require different approaches to perceive and interpret the images.

While radiologists are experiencing the transition from 2D projected radiography to 3D image datasets, image research has also to face the question of how this change would likely challenge previous observations and derive comprehensive conclusions from radiologists' practice with images from new imaging modalities. Despite the knowledge we have obtained from eye tracking system on 2D radiographic image, it is likely that 3D images interpretation is more relying on the combination of the characteristics of 2D and 3D images, which including human-computer interaction, 3D rendering presentation and navigation in third dimension. It is more important that we can understand the impact of new imaging modalities and image formats on radiologists' interpretation and therefore help radiologists to adapt and develop more efficient interpretation methods to improve their performance.

To the best of our knowledge there are very few, if any, researches in navigation and search patterns for medical 3D image dataset interpretation. Nevertheless, considerable efforts have been devoted to developing or designing 3D rendering and display methods to make image presentation more effective for medical detection diagnosis [13, 14]. Surface rendering, for example, is commonly used rendering method for displaying external structures and object shapes [15, 16]. Volumetric rendering methods, on the other hand, are more diagnostic relevant for revealing internal anatomical structures [17, 18]. One of the most commonly used

volumetric rendering methods is maximum intensity projection (MIP), which maximizes contrast on a rendered image by taking brightest voxel on a projected voxel ray [19]. Researchers have also tried to combine different rendering algorithms into one application to manage different anatomical structures and different diagnostic purposes. Various display workstations and user interface have been developed in order to achieve better image perception, ease reader-computer interaction and improve the efficacy of interpretation process [20, 21, 22, 23, 24, 25]. Our stereo display tested in this study is one of those attempts to improve radiologists' performance in lesion detection and classification. [22, 23]

All the works on 3D image data manipulation and presentation have significantly facilitated medical 3D data visualization. It is, however, unclear how well radiologists would adapt to the new technology and, more importantly, what kind of features or functionalities will likely improve radiologists' performance and maximize the utility value of the new technology. To understand radiologists' search pattern during interpretation of 3D image datasets, we have collected navigation data from a pilot study designed for ROC (Receiver Operating Characteristic)-type analysis for lung nodule detection on CT images and characterized the patterns that are related to the nodule detection and classification. The search patterns obtained from different display modes provided useful information on 3D image interpretation and possible improvement of display design.

Materials and Methods

A pilot study of lung nodule detection and classification on CT images was used for studying interpretation and navigation patterns. The detection and classification task was performed on three display modes (conventional slice-by-slice display, orthogonal MIP display and stereoscopic display), and radiologists' search patterns were collected and the performance was compared between the three display modes.

Data specification

Low dose lung CT images for lung cancer screening were acquired from multislice CT scanner (LightSpeed, GE, Milwakee), at a reconstructed thickness of 2.5-mm per slice and pixel resolution ranged from 0.69×0.69 mm^2 to 0.94×0.94 mm^2. There are about 100 axial images for each case, and a total of 30 cases were randomly selected from the lung cancer screening cases.

We have recruited six experienced staffed radiologists and two fellow radiologists to interpret the images. The primary task of interpretation was to detect and then classify any nodules equal to or larger than 3 mm in diameter with three distinct computer display modes, which are described in followings.

Image rendering

The display modes used in this study included slice-by-slice, orthogonal MIP rendering and stereoscopic view. Raw CT images were first processed with the convolution kernel provided by GE standard reconstruction software to form reconstructed images that are optimal for viewing lung tissues. The reconstructed images were then rendered based on the specification of each display mode. All renderings were precalculated and stored on hard disk for real-time display.

Slice-by-slice -- This is the most common display method adopted by radiologists for CT image interpretation. As images are read one at a time in sequence, no further rendering process was applied after the raw images were reconstructed with the lung kernel filtration. This set of single images was also included as a subset in the next two display modes.

Orthogonal MIP -- A stack of various number of CT slices were used to form MIP images. In this study, we implemented MIP images at thickness of 3, 5, 7, 9, 13 and 15 CT slices, respectively. Thickness of single slice was included in this display mode. For a given number of CT slices (thickness), a serial MIP images were rendered along the axial direction.

Voxel resampling was performed at axial direction (z-direction) to approximate isotropic voxel before performing 3D rendering. For orthogonal MIP rendering, the maximum value on an orthogonally projected voxel ray was selected as the final display value for each pixel on the MIP image.

Stereo perspective projection -- Slab thickness selection and voxel resampling used in orthogonal MIP rendering were also applied for stereo rendering. Linear perspective projection was applied to a stack of images to form horizontally shifted transformations of left- and right-eye images. Interocular distance (6.5-cm) and viewing distance between a viewer and computer screen (45-cm) were used to determine the angles of both eyes for perspective transformations.

Two rendering methods were employed for the stereo images [22, 23]. One was distance-weighted MIP rendering to produce high contrast images for nodule detection; and the other was distance-weighted averaging rendering to produce images of highly preserved local geometry for nodule classification. Distance-weighted algorithm incorporated in the stereo renderings provided transparency mechanism to adjust light transmission according to voxel locations. The detailed methods were included in references 22 and 23.

Display interface

A desktop personal computer was used for three display modes to display lung CT images. The computer has a central processor of 2.0 GHz ADM Athlon 64 3200+, 512 MB RAM, and a 128 MB NVIDIA Quadro FX 1100 graphics card. During stereo display, stereo effect was achieved through a shutterglasses (Stereo3D) controlled by frame-swap signals of displaying left-eye and right-eye images on the graphics card. A 21.0" (20.0" viewable) PerfectFlat CRT monitor, ViewSonic® Graphics Series G220f, was used in the display workstation. The monitor refresh rate was set to 144 Hz to produce stereo view without flickering effect.

A user interface was implemented using Microsoft Visual C++ API combined with OpenGL for image display and user interaction tools. Interactive operations during case interpretation basically involved navigation/search activity for lung nodules by moving along the axial direction throughout the lung area, and nodule assessment for any detected nodules. All the navigation/search related activities were conducted on a programmable keypad, which was dedicated to the specific needs for this study. The function keys on the programmable keypad can be used for selecting image axial viewing position and viewing volume (slab thickness), changing window/level settings, switching between MIP rendering and averaging rendering during stereo display, and toggling detected nodules.

An onscreen scoring form was designed and implemented for lung nodule classification. When radiologist clicks on a detected nodule, the scoring form with questionnaire related to the detected nodule would pop up for nodule assessment. We have also implemented mouse cursor as an onscreen ruler that can be used for nodule size estimation.

Study design

This study was designed for Free-response Receiver Operating Characteristic (FROC) type of detection. The task of the study was to detect any nodule-like feature and characterize it.

Randomization of case order and display modes was applied to each interpretation session to avoid bias caused by predictability from case order or particular mode. To avoid bias of familiarity, there were at least 14 days apart between two studies of same case with different display modes.

Data collection

We have recorded navigation patterns from four participating radiologists randomly selected to anonymize attributes associated with each individual. Navigation pattern during interpretation was collected by recording viewing volume (slab thickness) and viewing position at a 250 millisecond interval. For a detected nodule, the position of x, y, and z dimensions were recoded along with the other parameters, such as the likelihood probability of a nodule, the likelihood probability of malignancy, nodule shape, calcification and nodule size.

Data analysis

Interpretation time for each case was computed and compared between the modes. The navigation and nodule detection patters were visually analyzed and compared between the modes. Viewing volumes were analyzed from slab thickness recorded during case interpretation.

The nodule detection performance was determined by FROC analysis using JAFROC software (JAFROC, Chakraborty and Berbaum, http://www.devchakraborty.com/). The Figures of Merit from FROC were presented on a per-nodule basis. To verify the nodules, we used consensus results as the truth profile. The nodule-like features pooled from eight radiologists' interpretation in the three display modes were reviewed and verified by an experienced chest radiologist, who did not participate the study but had read and discussed the cases with other radiologists multiple times.

Results

Performance of nodule detection

The performance was evaluated based on consensus results of nodule location and likelihood probability. Total of 174 nodule-like features at the size of equal to or larger than 3-mm in diameter have been found in the 30 cases and 91 of them are true nodules. FROC analysis suggests that the stereo display resulted the performance that was better than the orthogonal MIP display, but was equivalent to the slice-based display, although no statistically significant difference was shown between the tree display modes. The Figures of Merits from the JAFROC software were, 0.57 (stereo display), 0.56 (slice-by-slice display) and 0.52 (orthogonal MIP display) for 8 radiologists, and 0.59 (stereo display), 0.61 (slice-by-slice) and 0.53 (orthogonal MIP display) for 4 radiologists whose navigation courses were recorded.

One of the efficiency measurements is interpretation time on each tested display mode. By averaging the time over 4 radiologists on each display mode, we have shown that the average interpretation time was significantly less with the stereo display (3.5 minutes) than with the slice-by-slice display (4.5 minutes), but was not much difference between the stereo display and the orthogonal MIP display (3.7 minutes).

Navigation pattern

The average viewing volume for the 3D displays was between 3 and 5 CT slices. There was no apparent difference in the preference of viewing volume between the stereo display and the orthogonal MIP display. When a region or a feature was in suspicious, a quick back-and-forth navigating across several slices was observed. This distinctive navigation pattern was more typically seen in the slice-by-slice display mode and in the nodules described as non-solid or semi-solid features. To interpret the case, the radiologists typically navigated through the dataset axially between the top and the base of the lung several times. The average number of such navigation rounds for the stereo display, the orthogonal MIP display and the slice-by-slice display were 3.4±4.3, 4.3±2.2 and 3.5±1.7, respectively.

The learning curve related to the 3D displays (the stereo display and the orthogonal MIP display) was demonstrated by the comparison of the navigation patterns at the beginning and the end of this study in figures 1, 2 and 3. The data was recorded from four radiologists' interpretations on each display mode. Comparing to the search course at the end of the study, the navigation patterns and viewing volume with the stereo (figure 3) and the orthogonal MIP (figure 2) displays were more complicated and dynamic at the initial stage of the study. Toward the end of the study, the navigation patterns became much smoother and more stabilized in both the stereo display and the orthogonal MIP display. The navigation patterns from the slice-by-slice display were, however, more like random search manner than a learning process when comparing the navigation patterens at the beginning and the end of the study (figure 1). Since case order was randomized at each interpretation session for each radiologist, the navigation patterns between radiologists shown in figures 1, 2 and 3 were not taken from the same cases.

Characteristic of missed nodules

We have compared missed nodules at apical lung area as well as the area close to diaphragm between the three display modes. Since a nodule could be detected 8 times (8 participating radiologists) in each display mode, it would be more appropriate to use number of detections, instead of number of nodules, for comparison. The total number of detections in the apical area and diaphragm area should be 128 and 112, respectively. In the apical area, there was a higher missed detection rate either in the stereo display (55%) or the orthogonal MIP display (55%) than that in the slice-by-slice display (42%). However, the difference was not such obvious in the lung area close to diaphragm, in which the missed detection rates were 36% for the stereo display, 38% for the orthogonal MIP display and 35% for the slice-by-slice display.

Further analysis from the search patterns revealed that some of the missed nodules were actually received extra attention from radiologists despite of no report being filed. Typical search pattern of the area, where a missed nodule resides and radiologist paid extra attention, are shown in figure 4A and 4B. There were about 25% of missed detections that received extra attention in the slice-by-slice mode, 15% in the orthogonal MIP mode and 16% in the stereo mode.

Structural characteristics of false detections

Most of falsely claimed nodules were various forms of scar tissues and vessels (table 1), in which scar tissues occurred more than vessels. Other structures that falsely recognized as nodules include bronchiectasis, atelectasis, and soft tissues. In the vessel group, more false detections were found in the orthogonal MIP display mode than in the slice based or the stereo display mode as shown in table 1.

Discussion

When medical imaging is rapidly evolving from 2D radiography to volumetric datasets, information presentation is also being changed. The main difference between 2D data and 3D data is that spacial information is not compressed in the 3rd dimension and therefore each image within a volume shares partial information that is much less than information in a projected radiographic image. To help radiologists more efficiently handle the volumetric data such as data from CT and MR, many programs were implemented to render and display the data to be visually comprehensible. The results and feedbacks seemed very positive regarding to the performance [13, 16,

26]. The actual clinical utility and impact, however, are not well documented and demonstrated. It is our main interest in this paper to present our preliminary results of interpretation patterns with volumetric datasets and to understand the impact of different display schemes on the search characteristics, for lung nodule detection and classification.

Even though the stereo display resulted generally less interpretation time and less falsely claimed nodules among the three tested display modes, the overall performance from the stereo display did not surpass the one from the slice-by-slice display. Subjective opinions and objective observations suggest that training effects significantly influence radiologists' search behavior and interpretation results. Of the three display modes tested in the pilot study, the stereo display has never been used or tried by the participating radiologists and the orthogonal MIP display has been experienced to a very limited extent. We observed from the navigation patterns that, at the beginning of the study, radiologists were vigorously tuning their search patterns to try to find the optimal search pattern and optimal viewing volume with the two 3D displays, suggesting active involvement of learning and improving, which could including familiarization to the 3D displays and optimization of search strategy. In contrast, the navigation patterns in the slice-by-slice display were more randomized and undifferentiated between the beginning and the end of the study (figure 1). Further, we observed that when interpreting cases with the 3D displays, radiologists tended to adjust the viewing volume from initial thick slab to single slice during early stage of the study. As single slice was the subset of the viewing volume in these volumetric display modes, the preference for the single slice suggested strong influence of training effect to radiologists' interpretation behavior.

Evidence of training effect further comes from the observation that there were more attention-paid missed nodules associated with the slice-by-slice display than either with the orthogonal MIP or with the stereo display. Radiologists have been trained conventionally in single projected radiographic image interpretation and maintained consistent practice manner for scrutinizing this kind of images carefully. But they are not extensively exposed to volumetric display, and meantime lacking of systematical training for volumetric data interpretation. Furthermore, since volumetric display can show more information in one view and clear geometrical relationship for easy understanding compared to single slice based display, radiologists may be over-confident for their observation and tend to neglect some subtle structures needed for more attention and/or different skills in 3D view. Appropriate training and practice, therefore, is necessary for achieving optimal performance with 3D display device and new display technology.

While novelty seemed to substantially affect navigation patterns and the performance, other factors associated with our 3D displays may also influence the results. Despite similarity in the navigation patterns and in the use of thickness information, the orthogonal MIP rendering and the stereo view showed some differences in nodule detection. Vessel-like structures were much easier to be mistakenly recognized as nodules in the orthogonal MIP display as compared to that in the stereo display. Overall, the orthogonal MIP resulted more false positive findings than stereo display (table 1) and the lowest performance score among the three display modes, although with no statistically significant difference. The low performance and high false positive rate of the orthogonal MIP rendering are most likely attributed to superimposed structures of monoscopic thick slab. Despite high contrast volumetric images, orthogonal MIP rendering may not produce correct geometric representation of volumetric objects due to that the algorithm takes the highest intensity along each projection ray, which may very well not preserve structural continuity between adjacent pixels in the rendered image. The stereoscopic rendering, on the other hand, was implemented with perspective transformation and transparency mechanism so that superimposition was not introduced and local geometric information was better preserved, especially with averaging method.

When lung nodules are neighbored with similar intensity of non-lung tissues in a thick viewing volume, they are likely to be missed due to camouflage effect. We have examined two places where lung tissue could be obscured by surrounding structures. One was apical lung area, where lung tissues are closely surrounded by rib cage. The other one was the area close to diaphragm. The results indicate that there were more missed detections with either the stereo or the orthogonal MIP display than with the slice-by-slice display in the apical area, while no such difference showed in the diaphragm area between the three displays. As obscuration can lower the conspicuity of the nodules, other factors, such as structure density and shape relationship, may also have effect on the detection as suggested by the different results from the two areas.

In this study, we have not implement multiple reformations for different viewing angles because of the complexity of preparing prestaged multiple reformations. Results from other researches and our current project of real-time rendering on programmable graphics units indicate that volumetric displays that allow multiple reformatted viewing angles by rotating images can help reduce ambiguity caused by some poorly differentiated spacial relationships including tissue superimposition [26, 27, 28]. The advantage of multiple reformations can be more appreciated by volumetric displays than single slice based display. Multiple views for single slice are geometrically discontinuously transformed because they lack the information of the third dimension and require intensive mental work on geometrical correlations between two viewing angles. When viewing volumetric data, volume can be smoothly transformed between two viewing angles by rotating objects in 3D space to produce natural continuation of views of the objects. The improvement of structure differentiation may be further enhanced by making non-interested tissue transparent to reduce the ambiguity.

Although more 3D imaging modalities are being employed for medical screening and diagnosis, slice-by-slice display is still predominantly being used as a primary viewing method for interpretation. Adopting volumetric displays, therefore, involves learning process that extents and transforms current 2D understanding of medical images to the knowledge of volumetric information discovery. Effective utilization of 3D display for medical volumetric data relies both on software design and user training. Our preliminary data from a pilot study for lung nodule detection on CT images indicate that current 3D displays can be further improved by understanding radiologists' interpretation behavior and diagnostic performance. In addition, stereoscopic display produced more efficient interpretations and lower false position detections comparing to other displays, and has promising potential for improving radiologists' performance and efficiency of 3D dataset interpretation.

Acknowledgement

This work is sponsored in part by the US Army Medical Research Acquisition Center, 820 Chandler Street, Fort Detrick, MD 21702-5014 under Contract PR043488, and also by grant CA80836 from the National Cancer Institute, National Institutes of Health. The content of the contained information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

References

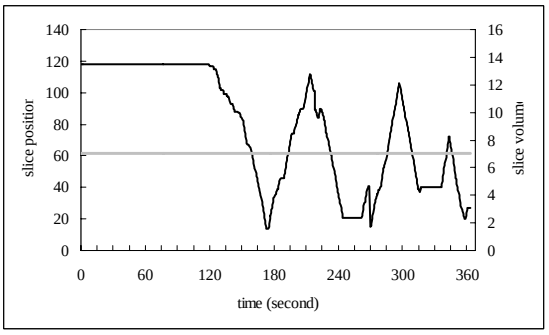
1. Krupinski EA. Visual search of mammographic images: influence of lesion subtlety. *Acad Radiol*. 2005; 12:965-969

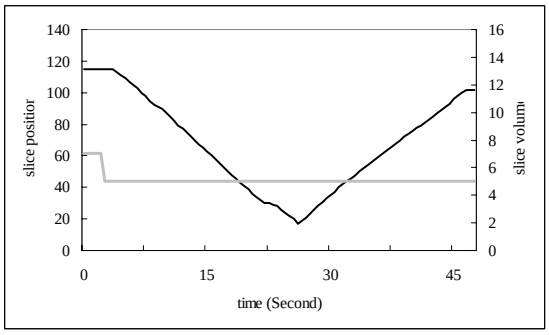
2. Krupinski EA, Berger WG, Dallas WJ, et al. Searching for nodules: what features attract attention and influence detection? *Acad Radiol*. 2003; 10:861-868
3. Nodine CF, Mello-Thoms C, Kundel HL, at el. Time course of perception and decision making during mammographic interpretation. *AJR Am J Roentgenol*. 2002; 179:917-923
4. Kundel HL, Nodine CF, Toto L. Searching for lung nodules. The guidance of visual scanning. *Invest Radiol*. 1991; 26:777-781
5. Kundel HL. Visual cues in the interpretation of medical images. *J Clin Neurophysiol*. 1990; 7:472-483
6. Kundel HL, Nodine CF, Krupinski EA. Searching for lung nodules. Visual dwell indicates locations of false-positive and false-negative decisions. *Invest Radiol*. 1989; 24:472-478
7. Kundel HL, Nodine CF, Thickman D, et al. Searching for lung nodules. A comparison of human performance with random and systematic scanning models. *Invest Radiol*. 1987; 22:417-422
8. Krupinski EA, Roehrig H. The influence of a perceptually linearized display on observer performance and visual search. *Acad Radiol*. 2000; 7:8-13
9. Nodine CF, Kundel HL, Lauver SC, et al. Nature of expertise in searching mammograms for breast masses. *Acad Radiol*. 1996; 3:1000-1006
10. Barish MA, Rocha TC. Multislice CT colonography: current status and limitations. *Radiol Clin North Am*. 2005; 43:1049-1062
11. Rowe VL, Tucker SW Jr. Advances in vascular imaging. *Surg Clin North Am*. 2004; 84:1189-1202
12. Israel GM, Bosniak MA. Renal imaging for diagnosis and staging of renal cell carcinoma. *Urol Clin North Am*. 2003; 30:499-514
13. Aufort S, Charra L, Lesnik A, et al. Multidetector CT of bowel obstruction: value of post-processing. *Eur Radiol*. 2005; 15:2323-2329
14. Remy J, Remy-Jardin M, Artaud D, et al. Multiplanar and three-dimensional reconstruction techniques in CT: impact on chest diseases. *Eur Radiol*. 1998; 8:335-351
15. Bomans M, Hohne KH, Tiede U, et al. 3-D segmentation of MR images of the head for 3-D display. *IEEE Transactions on Medical Imaging*. 1990; 2:177-183
16. Hoehne TU, Bomans KH, Pommert M, et al. Investigation of medical 3D-rendering algorithms. *IEEE Comp Graphics & Applications*. 1990; 2:41-53
17. Calhoun PS, Kuszyk BS, Heath DG, et al. Three-dimensional Volume Rendering of Spiral CT Data: Theory and Method. *Radiographics*. 1999; 19:745-764
18. Kuszyk BS, Heath DG, Ney DR, et al. CT Angiography with Volume Rendering: Imaging Findings. *AJR, American Journal of Roentgenology*, 1995; 165:445-448
19. Fishman EK, Ney DR, Heath DG, et al. Volume Rendering versus Maximum Intensity Projection in CT Angiography: What Works Best, When, and Why. *RadioGraphics* 2006; 26:905-922
20. Demiralp C, Jackson CD, Karelitz DB, Zhang S, Laidlaw DH. CAVE and fishtank virtual-reality displays: a qualitative and quantitative comparison. IEEE Trans Vis Comput Graph. 2006 May-Jun;12(3):323-30.
21. Levin D, Aladl U, Germano G, et al. Techniques for efficient, real-time, 3D visualization of multi-modality cardiac data using consumer graphics hardware. *Comput Med Imaging Graph*. 2005; 29:463-475
22. Wang XH, Maitz GS, Leader JK, et al. Real-time stereographic display of volumetric datasets in radiology. 2006, SPIE, Electronic Imaging vol 6055, 1A-1 - 1A-6.
23. Wang XH, Walter FG, Fuhrman CR, et al. Stereo CT Image Compositing Methods for Lung Nodule Detection and Characterization. Academic Radiology 2005, 12:1512-1520.
24. Yee DK, Lee W, Kim D, et al. RadGSP: a medical image display and user interface for UWGSP3. *Proc. SPIE* 1991; 1444:292-305
25. Fuchs H, Levoy M, Pizer SM. Interactive visualization of 3D medical data. *Computer*. 1989; 22:46-51
26. Salvolini L, Bichi Secchi E, Costarelli L, et al. Clinical applications of 2D and 3D CT imaging of the airways--a review. *Eur J Radiol*. 2000; 34:9-25
27. Ou P, Celermajer DS, Calcagni G, et al. Three-Dimensional CT Scanning: A Novel Diagnostic Modality in Congenital Heart Disease. *Heart*. 2006; Epub ahead of print: http://heart.bmjjournals.com/cgi/content/abstract/hrt.2006.101352v1
28. Durkee NJ, Jacobson J, Jamadar D, et al. Classification of common acetabular fractures: radiographic and CT appearances. *AJR Am J Roentgenol*. 2006; 187:915-925

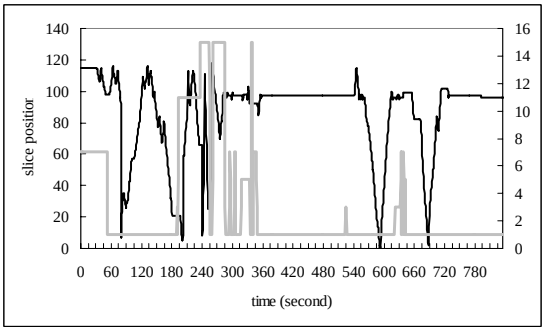
Table 1. Distribution of false positive findings in different structural groups.

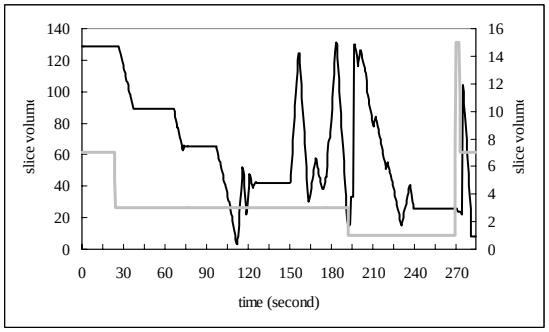
	Stereo	Orthogonal MIP	Slice by slice	Total
Vessel	11	27	18	56 (32%)
Scar	23	32	33	88 (51%)
Other	10	7	12	29 (17%)
Total	44 (26%)	66 (38%)	63 (36%)	173 (100)

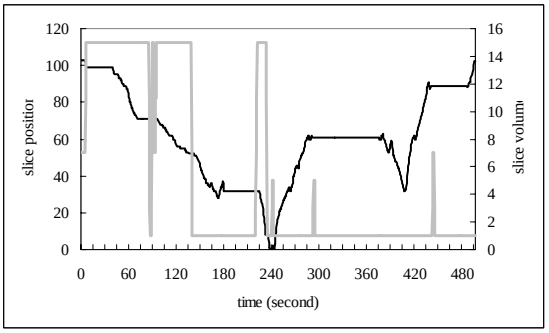
Figure 1. Navigation patterns from four radiologists in **stereo mode**. Each graph is the navigation recorded from one case interpretation. Two graphs in each row are taken from one radiologist's interpretations in the beginning of the study (left) and the end of the study (right). The dark solid line represents viewing positions referenced to the scale on left axis and grey solid line represents viewing thickness referenced to the scale on right axis, with time.

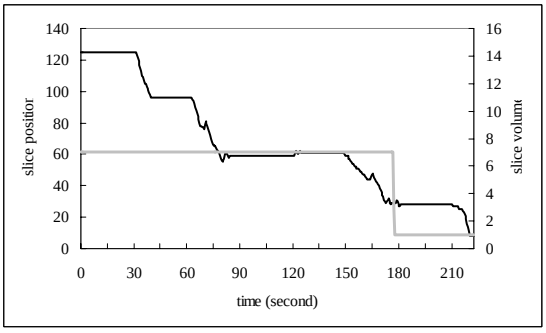


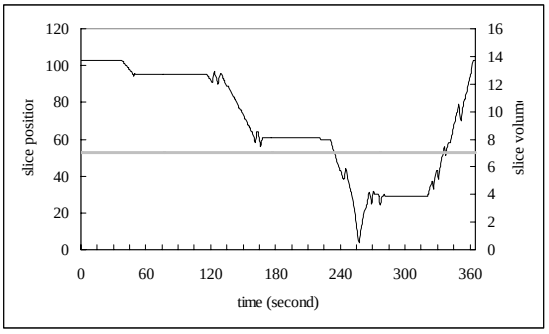












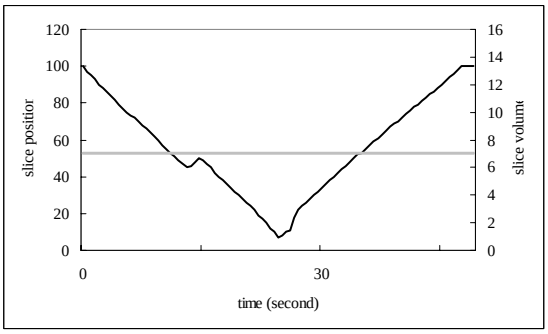


Figure 2. Navigation patterns from four radiologists in **orthogonal MIP mode**. Each graph is the navigation recorded from one case interpretation. Two graphs in each row are taken from one radiologist's interpretations in the beginning of the study (left) and the end of the study (right). The dark solid line represents viewing positions referenced to the scale on left axis and grey solid line represents viewing thickness referenced to the scale on right axis, with time.

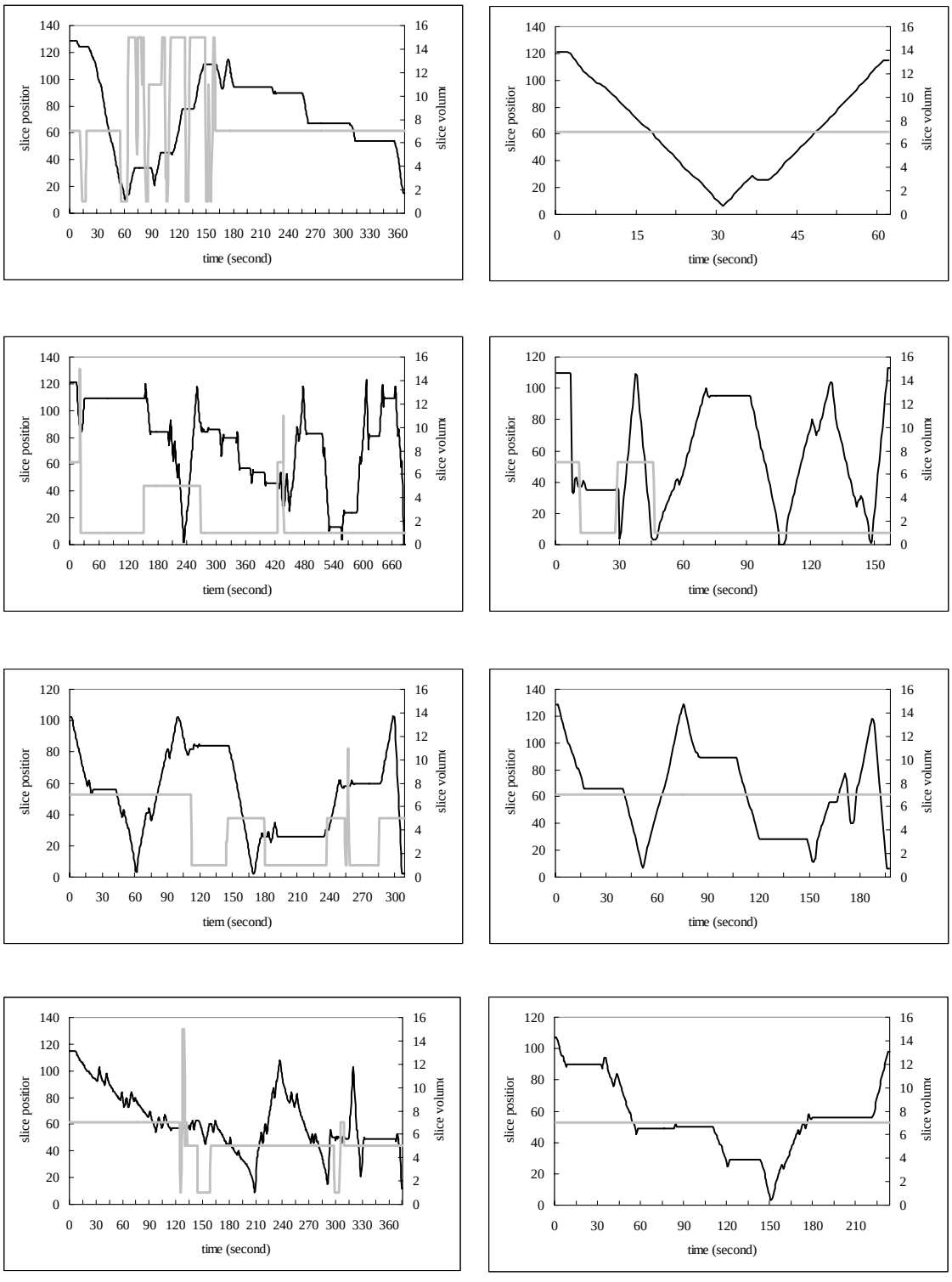


Figure 3. Navigation patterns from four radiologists in **slice-by-slice mode**. Each graph is the navigation recorded from one case interpretation. Two graphs in each row are taken from one radiologist's interpretations in the beginning of the study (left) and the end of the study (right). The dark solid line represents viewing positions with the time.

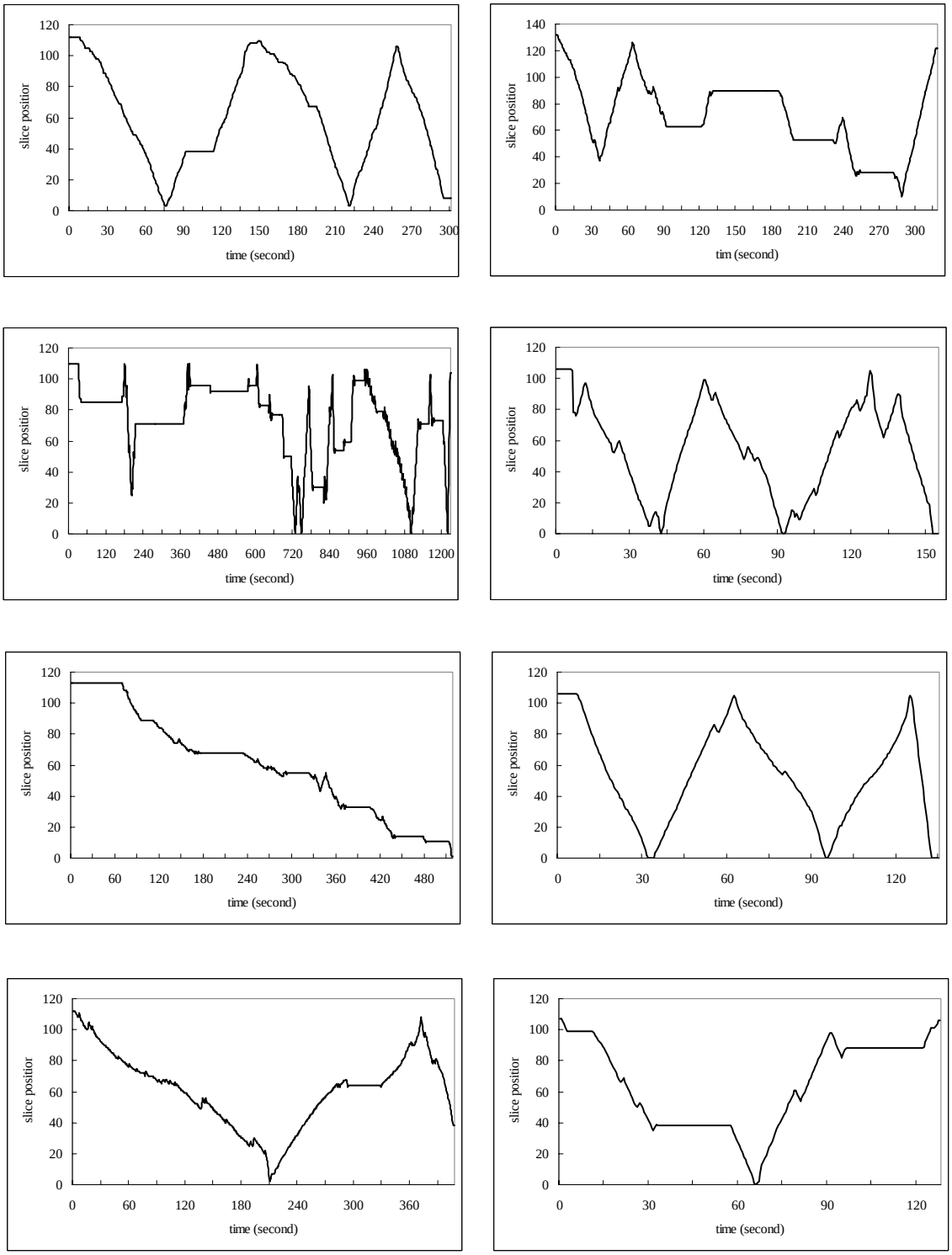
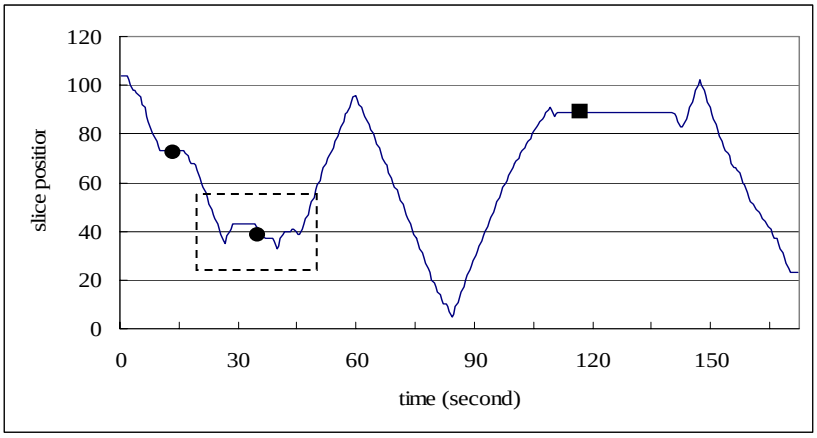


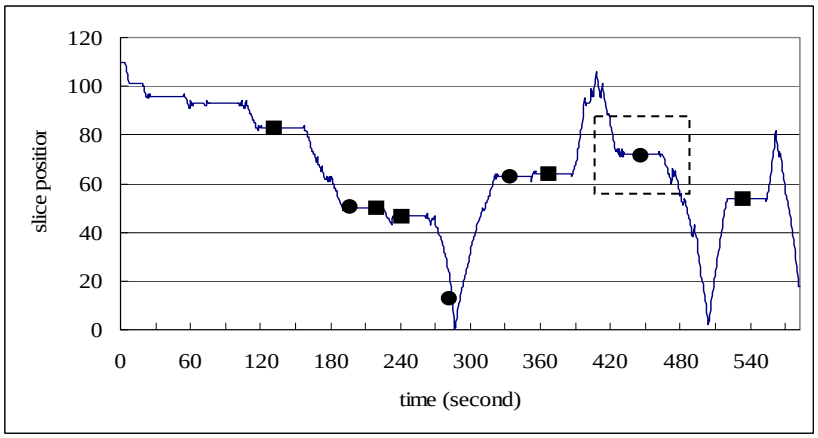
Figure 4. The diagrams A and B illustrate the missed nodules that received extra attention.

— viewing position ⬚ location of missed nodules ● true nodule ■ false nodule

A.



B.



Appendix C

Real time stereographic rendering and display of medical images with programmable GPUs

Xiao Hui Wang and Walter F. Good

Department of Radiology, University of Pittsburgh, School of Medicine, Pittsburgh, PA

Abstract

The amount of volumetric data being acquired in radiology is rapidly increasing. To maintain performance and efficiency in reading this data, it is desirable to be able to display the data as 3-D monoscopic- or stereoscopic-renderings, with real-time interactive control by radiologists. This paradigm has not been widely adopted because of the difficulty and expense of providing the required computational resources. With the availability of newer commodity graphics processing units (GPUs) for personal computers, it may be possible to overcome the computational impediments to interactive 3-D displays. This study compared the frame rates that can be achieved on CPUs to those that can be achieved by exploiting GPUs, and finds that GPUs are capable of rendering large 3-D datasets at real-time interactive rates.

Introduction

Inherently 3-D medical imaging modalities, such as Computerized Tomography (CT) and Magnetic Resonance (MR) imaging systems, are generating an ever increasing volume of image data that must be reviewed by radiologists. This trend will almost certainly continue into the future as radiologists, in an effort to increase spatial resolution, depict 3-D volumes by using thinner, but more numerous, slices.

Current Display Paradigms -- By far, the most common method used for viewing inherently 3-D data has been reading 2-D slices sequentially, from the 3-D dataset, in a slice-by-slice mode, a laborious and error prone process, or viewing the data as projections of thicker sections comprised of multiple adjacent slices.

It is known that the visibility of certain kinds of subtle features can be increased by presenting the data as a thicker 3-D volume [1,2], rendered with an appropriate projection algorithm. This is normally achieved by combining the thin slices directly acquired during volumetric imaging to form a thicker slab, and then projecting this slab onto a 2-D display, but increasing the thickness of projected volumes can cause ambiguities due to the superposition of tissues. Also, use of an averaging process to combine slices can reduce the contrast of features that are small relative to the thickness of the resulting slab. As slabs become thicker by adding more of the originally acquired thin slices, the contrast of smaller features, which often are visible on only one or two thin slices, may be reduced by averaging with the remaining thin slices [1]. As slices become thinner, the signal-to-noise ratio in individual slices is reduced making it more difficult to detect certain kinds of features, and at the same time, the number of slices that must be read increases. Furthermore, the process of reading individual slices sequentially forces viewers to reconstruct mentally the 3-dimensional structure, and does not permit the reader's visual system to take full advantage of correlations between adjacent slices to improve apparent signal-to-noise ratios.

Various methods for 3-D display of volumetric radiographic datasets have been devised to make the reading process more efficient, but they have not been widely adopted because of certain performance limitations. Specifically, the task of rendering 3-D datasets in a form that is suitable for radiological applications is computationally intensive and it has not been possible to perform these calculations sufficiently fast to be able to provide radiologists with real-time interactive displays, except on superpremium computers. There is a consensus that, without real-time interactivity, volumetric display (monoscopic or stereoscopic) is often not justified by the added complexity.

Potential Role of Stereographic Displays – Stereographic display of 3-D radiographic datasets, which takes full advantage of readers' binocular vision, may provide benefits beyond those attributed to monoscopic 3-D display [3]. Certain kinds of objects can be detected in a stereo 3-D display of data, which cannot be detected when the data is viewed in a slice-by-slice manner. Stereo projection can improve the visibility of objects by enhancing features that are correlated between slices, while reducing noise in a manner analogous to the signal-to-noise improvements obtained by averaging slices or MIPs – but stereo projection does not introduce tissue superposition ambiguities that would be caused by these methods [4]. Nevertheless, stereographic presentation has received even less attention than monoscopic 3-D because it further increases the computational burden.

Application of GPUs – With the evolution of commodity graphics processing units (GPUs) for accelerating games on personal computers, over the past couple of years, the amount of computing power that is available for rendering complex scenes has been rapidly increasing. GPUs may be capable of performing a wide range of reconstruction, volume reformatting and stereo projection in real-time under user control. In particular, the most recent GPUs are approaching a performance level where real-time interactivity with stereographic displays is feasible.

GPUs are organized as pipelined parallel processors. They differ from general purpose processors, that basically perform one instruction at a time and need to have the result returned immediately, in that they process parallel streams of independent data and can wait for an individual result as long as the entire dataset is processed quickly [5]. In this sense, they are ideal for tasks that are computationally intensive in volumetric rendering of 3-D datasets. Dietrich, et al, report that they were able to achieve real-time rendering of a 512 x 512 x 100 liver CT dataset on a 2 GHz Pentium 4, with a ATI 9800 GPU, though they were primarily concerned with only the volume clipping component of the rendering algorithm [6].

Several researchers, including our own, have shown the potential benefit of GPUs for efficient image manipulation and visualization within medical applications [7-12]. For example, Briggs, et al, have demonstrated a display for volumetric electrical impedance tomography [7]. While their datasets are smaller than many that occur in radiology, they were able to achieve real-time performance. A Doppler-ultrasound display was implemented by Heid, et al, by

exploiting the performance of a GPU [8]. GPU-based programming has been implemented for interactive 4-D motion segmentation and volume rendering of cardiac data and has resulted efficient data processing and visualizing with high quality and at real-time speeds [9]. GPUs were also demonstrated to be efficient in generating high quality reconstructed radiographs from portal images and CT volume data for radiation therapy [10]. Sorensen, et al, have also applied the technique to surgical simulation of the liver to achieve a real-time performance, where surface rendering involving dynamic geometric transformations and texture manipulations were implemented on a GPU [11]. These specialized applications can often achieve significant levels of performance by optimizing their systems for the application, but these systems do not necessarily retain that performance when used in a different context.

We have tested the feasibility and efficacy of performing renderings on GPUs for stereo display of medical 3-D dataset. Previously, we prestaged and prerendered stereo pair renderings of lung CT images for display. Because of different viewing positions and viewing volumes, rendering a complete set of image pairs for a case took a substantial amount of time and consumed vast storage space. Such a practice may work within certain research environments, but is not practical for the general clinical settings, where real-time rendering and manipulation are necessary for prompt and accurate diagnosis.

While GPUs have been applied to a number of radiological imaging tasks, their potential performance characteristics are not well understood. This study is an attempt to measure frame rates that can be achieved for stereographic rendering on a GPU and compares these to rates that can be achieved on CPUs alone.

Methods

Data set – Images used for developing GPU-based rendering and display were obtained from a 4-detector CT scanner (LightSpeed Plus, GE medical Systems, Milwaukee, WI) for lung cancer screening program. The CT images were acquired in the axial plane and reconstructed to a thickness of 2.5 mm/slice with lung kernel reconstruction algorithm provided by GE standard software. The pixel size on each slice ranges from 0.63-mm × 0.63-mm to 0.92-mm × 0.92-mm. There are approximately 512×512×100 data voxels for a typical lung CT case in our dataset.

Hardware -- The study was run on an off-the-shelf personal computer with a 2.0 GHz AMD Athlon 64 3200+ processor and 512 MB of RAM. The computer is equipped with a 128 MB NVIDIA Quadro FX 1100 graphics card, which has build-in support for stereographic buffering system to hold left- and right-eye images in separate frame buffers and to swap frame buffers for a frame-swapped display. The stereo image pairs are viewed either on CRT monitors via shutterglasses controlled by frame-swapping signals or on superimposed cross-polarized displays via passive polarizing eyeglasses.

Volume rendering for stereo display – Two rendering methods, Maximum Intensity Projection (MIP) and averaging, have been implemented to generate stereo pairs of the lung CT images. Because lesions must be detected before they can be evaluated, high contrast MIP images were preferable for lesion detection while images rendered by averaging, which preserves local geometry, were preferable for lesion evaluation [12,13]. The rendering process for both MIP and averaging in this application involves perspective transformation [14], transparency modeling based on optical occlusion/distance characteristics, and ray casting [12-13,15-18]. All the rendering processes that we have previously performed on CPU card can be now processed on a programmable GPU card.

Rendering on GPUs – Stereographic compositing and display was implemented and compiled in the OpenGL and Cg languages on NVIDIA programmable GPUs. A flowchart, shown in Figure 1, illustrates the operations performed on GPU card.

For a given slab thickness, a vertex block with dimensions of 512×512×thickness was generated to include all vertices for perspective transformation and texture-coordinates. The dimensions of each vertex were approximated so as to be isotropic in all three axes (x, y and z) based on acquired x and y dimensions. Vertex-coordinates and texture-coordinates were then specified and interpolated during rasterization before being input to the vertex and fragment programs.

A sufficient number of interpolated slices were generated to provide continuity of display in the axial direction. Typically, for a dataset such as the one employed in this project, 3 interpolated slices are generated for every real slice.

Perspective projection in ray casting was performed in vertex program for each input vertex. The matrices for perspective transformation were determined by a presetting of eye-offsets and viewing distance. In the case of stereo compositing, the projection centers for the left- and right-eye images are offset laterally relative to each other. The parallax value for each eye-offset is set close to 1° to achieve stereo depth perception while avoiding excessive eyestrain. The rotation transform was also performed in vertex program. Transformed vertices that were out of the clip volume were not used for display. An example of Cg vertex program for vertex transformations is shown in Code 1.

Code 1.

```
vertOutput main (       float4 Position : POSITION,
                        float4 Texcoord : TEXCOORD0,
                        uniform float4x4 rotate_x,
                        uniform float4x4 rotate_y,
                        uniform float4x4 translate_matrix,
                        uniform float4x4 perspective_matrix
                )
{
        vertOutput OUT;
        float4 rot_xP, rot_yP, tPosition, pPosition;
        rot_xP=mul(rotate_x,Position);
        rot_yP=mul(rotate_y,rot_xP);
        tPosition=mul(translate_matrix,rot_yP);
        pPosition=mul(perspective_matrix, tPosition);
        OUT.Position=pPosition;
        OUT.texcoord=Texcoord;
        return OUT;
}
```

Once a vertex has been geometrically transformed to a proper position, texture mapping for the vertex takes place in a fragment program. The 16-bit lung CT volume data (approximately 512×512×100) was loaded into the graphics memory to serve as a 3-D texture map. Texture values were automatically interpolated in the texture map with the OpenGL linear filter function for a given texture-coordinate. Occlusion/distance based transparency and window-level settings were also implemented in the fragment program. A Cg code fragment implementation is shown in Code 2.

Code 2.

```
float4 main (           vertOutput IN,
                        uniform sampler3-D testTexture,
                        uniform float window_level,
                        uniform float transparency_coef;
        ) : COLOR
{
        float4 color;
```

```
        float temp, tt;

        temp=tex3-D(testTexture, IN.texcoord);
        tt=(temp-window_level) * transparency_coef;
        color=tt;

        return color;
}
```

The final rendering process for displayed pixels was actualized by implementing the openGL blending functions. For MIP rendering the display value of each pixel was rendered by taking the maximum value among the points of a projection ray using MAX blending function, while for averaging rendering the display value was rendered by adding distance-weighted fractions of each fragment along a projection ray to the pixel using the ADD blending function.

Rendering on CPU card – The stereo image pairs for a lung CT case were prestaged and precalculated for all volume sizes between 1 up to 45 interpolated slices (see column 1 in table 1 and table 2) at all axial viewing positions. The detailed methods can be found in ref. In brief, we used trilinear interpolation to resample the data for a given volume of CT images to achieve final pixel dimension close to isotropic. Perspective transformation and ray casting based on compositing methods were performed for each pixel on stereo images. For MIP rendering the highest voxel value along a projection ray was used for projection value, while for averaging rendering each voxel value along a projection ray contributed a fraction to the final projection value.

Other functionality – An OpenGL based window display was built for displaying both GPU- and CPU-based stereo images. Specifically, window-level adjustment, viewing volume and viewing position selection and choice between MIP rendering and averaging rendering, were implemented. Image rotations were only performed by GPU-based rendering.

Results

The GPU-based program achieved real-time rendering and real-time display rates without any perceptible delay in the display of successive frames, following a user controlled frame switch command. We found no difference in frame rates between renderings by MIP and by averaging. A comparison of the rendering rates between GPU- and CPU-rendering, for our lung CT dataset, is shown in Tables 1 and 2. Table 1 lists the frame rate measurements of stereo compositing on GPU card as well as on CPU card at various volume sizes. The highest volume we rendered for lung CT images is 45 interpolated slices, which is about the thickness of 15 real CT slices at 2.5-mm. When we reviewed various stereo images with several experienced radiologists, we found that the preferred viewing volume for detection and diagnosis ranged from 3 to 7 real slices (i.e., 9 to 21 interpolated slices), and 15 slices (i.e., 45 interpolated slices) contained too much information to be useful for detection and diagnosis. Even with volume of 15 slices, which has more than 23 million vertex rendering processings (512×512×45×2 stereo images), we still achieved a rate at 5-frames per second. Rendering performed on the CPU card resulted in much slower frame rates and would not give the impression of real-time interactivity. If we precalculate all of these stereo image pairs for a case, it would take less than a minute on the GPU card versus more then 20 minutes on the CPU. Implementing rotation on the GPU card did not measurably reduce frame rates for the data volumes used in this study, as shown in Table 2.

Discussion

Traditionally, 3-D medical image datasets are rendered predominantly on CPUs to generate precalculated images that can be prestaged for reading by radiologists. This preprocessing procedure puts many constraints on the review process and, at the same time, consumes a substantial amount of storage space and CPU time. These CPU-based processes most likely will be replaced in the near future by processes performed on the advanced graphics cards, due to the fact that these cards are becoming readily available and their real-time processing speed and improved arithmetic precision is makes them suitable for the processing of many types of radiological images. The study presented in this paper shows that GPU-based rendering can achieve real-time interactive stereo display rates for lung CT images up to volumes larger than the optimal volume used for diagnosis. Monoscopic rendering rates, though not measured in this study, would likely be nearly double the stereoscopic rates for a given volume.

The benefit of using GPUs processing power can be widely appreciated in medical image detection and diagnosis. As show in Table 1, GPU-based programming renders stereo pairs in real-time for as many as 45 slices (more than 23 million vertices) and gives no perceptional delay between frame changes. The capability of real-time process eliminates the constraints from prestaged paradigms. Viewing angles, for example, can be important for detection and differentiation of an object. It is, however, impractical and impossible to prestage and precalculate all viewing angles for a set of images, or to perform smooth rotations. Whereas programmable GPUs perform real-time renderings, rotation functionality can be seamlessly and smoothly implemented during rendering process and consumes negligible GPU processing time compared to the overall processing time, as shown in Table 2.

From research conducted by others and our previous studies, we have observed that no single algorithm can meet all the requirements of clinical tasks. We have demonstrated that for stereo display, MIP rendering is the best for detection owing to the high contrast of rendered images, but not optimal for classification because of lack of local geometric fidelity in the rendered images. On the other hand, rendering by averaging will preserves local geometry despite providing low contrast of the rendered images. The two renderings can be used for different tasks during medical image interpretations. We have implemented this mechanism in CPU-based prestaged calculations and display, and the results were satisfactory at the expense of longer processing time and much more storage space. The GPU-based programming not only naturally solved this problem of dynamically switching between MIP and averaging renderings, but can also, in general, implement any algorithms, whichever needed, specific to the task in real time. This will dramatically improve efficacy of image presentation and diagnostic performance.

Acknowledgement

This work is sponsored in part by the US Army Medical Research Acquisition Center, 820 Chandler Street, Fort Detrick, MD 21702-5014 under Contract PR043488, and also by grant CA80836 from the National Cancer Institute, National Institutes of Health. The content of the contained information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

References

1 Brown DG, Riederer SJ. Contrast-to-noise ratios in maximum intensity projection images. Magn Reson Med 1992; 23:130-137.

2 Keller PJ, Drayer BP, Fram EK, Williams KD, Dumoulin CL, Souza SP. MR angiography with two-dimensional acquisition and three-dimensional display. Work in progress. Radiology 1989; 173(2):527-32.
3 Roberts JW, Slattery OT. Display characteristics and the impact on usability for stereo. *Proc. SPIE, Stereoscopic Displays and Virtual Reality Systems VIII* 2000, 3957:128-138.
4 Maidment ADA, Bakic PR, Albert M. Effects of quantum noise and binocular summation on dose requirements in stereomammogry. Med Phys 2003; 30(12):3061-3071.
5 Ujval J. Kapasi UJ, Rixner S, Dally WJ, Khailany B, Ahn JH, Mattson P, Owens JD. Programmable Stream Processors. Computer. August 2003; (Vol. 36, No. 8) 54-62.
6 Dietrich CA, Nedel LP, Olabarriaga SD, Comba JLD, Zanchet DJ, da Silva AMM, Montero EF. Real-time interactive visualization and manipulation of the volumetric data using GPU-based methods. SPIE Medical Imag 2004; 5367:181-191.
7 Briggs NM, Avis NJ, Kleinermann F. A real-time volumetric visualization system for electrical impedance tomography. Physiol. Meas. 21(2000)27-33.
8 Heid V, Evers H, Henn C, Glombitza G, Meinzer HP. Interactive real-time Doppler-ultrasound visualization of the heart. Stud Health Technol Inform. 2000;70:119-25.
9 Levin D, Aladl U, Germano G, Slomka P Techniques for efficient, real-time, 3-D visualization of multi-modality cardiac data using consumer graphics hardware. 2005, Computerized Medical Imaging and Graphics 29: 463-475.
10 Khamene A, Bloch P, Wein W, Svatos M, Sauer F Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy. 2006, Medical Image Analysis 10: 96-112.
11 Sørensen TS, Mosegaard J Haptic feedback for the GPU-based surgical simulator. 2006, Medicine Meets Virtual Reality 14:523-528.
12 Wang XH, Good WF, Fuhrman CR, Sumkin JH, Britton CA, Warfel TE, Gur D. Stereo Display for Chest CT. Proc SPIE EI2004 2004; In Press.
13 Wang XH, Good WF, Fuhrman CR, Sumkin JH, Britton CA, Warfel TE, Gur D. Projection Models for Stereo Display of Chest CT. Proc SPIE MI2004 2004; In Press.
14 Mortenson, ME. Computer Graphics Handbook: Geometry and Mathematics, 1990
15 Drebin RA, Carpenter L, Hanrahan P. Volume rendering. Comput Graph 1988; 22:65-74.
16 Levoy M. Display of surfaces from volume data. IEEE Comput Graph Applicat 1988; 8:29-37.
17 Ney DR, Drebin RA, Fishman EK, Magid D. Volumetric rendering of computed tomographic data: principles and techniques. IEEE Comput Graph Applicat 1990; 10:24-32.
18 Ney DR, Fishman EK, Magid D, Robertson DD, Kawashima A. Three-dimensional volumetric display of CT data: effect of scan parameters upon image quality. J Comput Assist Tomogr 1991; 15:875-885.

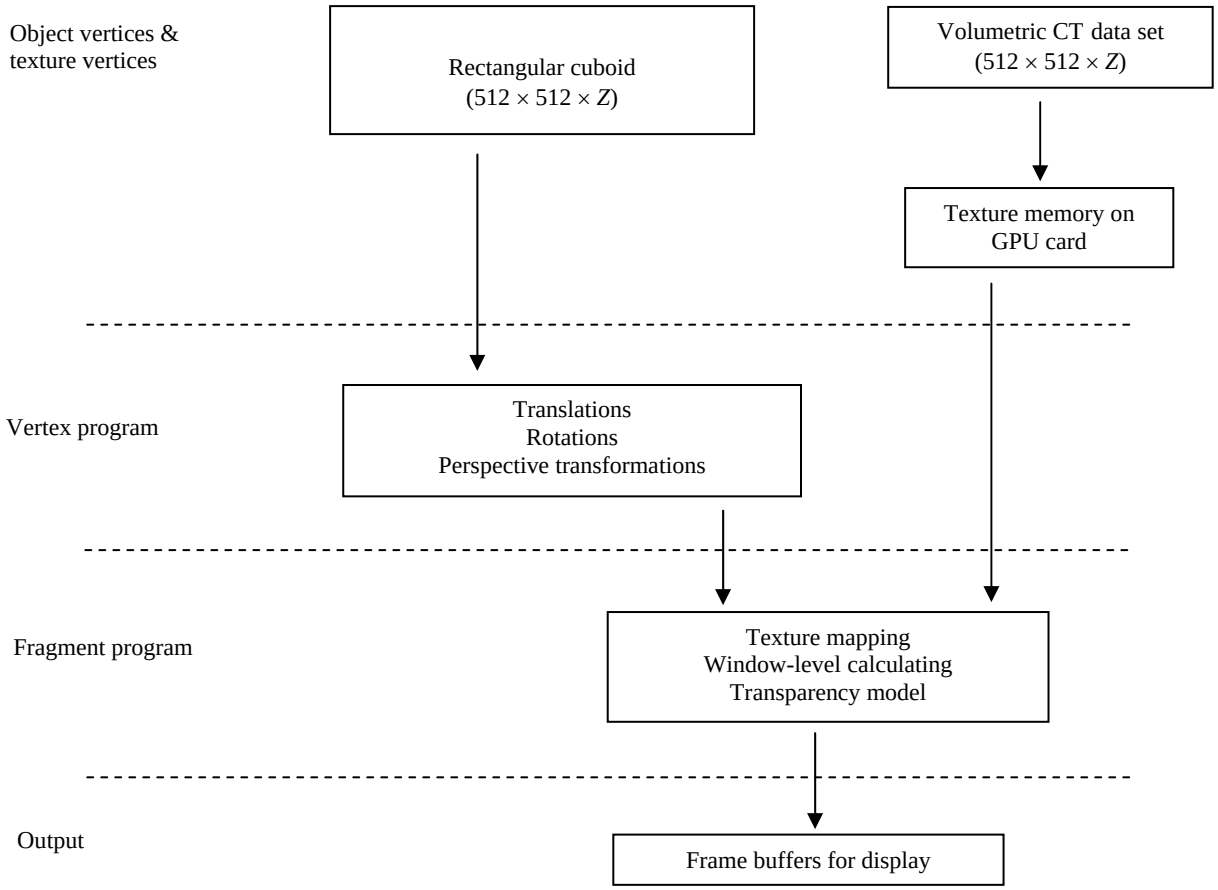


Figure 1. A diagram of stereo image rendering process on GPU card. *Z* is the depth measure of a given rendering volume.

Number of interpolated slices	GPU (stereo pairs per second)	CPU (stereo pairs per second)
1	103.3	-
9	20.1	1.3
15	13.2	0.8
21	10.1	0.5
33	6.6	0.3
45	5.0	0.2

Table 1. Frame rates measured as stereo pairs per second for rendering on GPU card and CPU card at different number of interpolated slices.

Number of interpolated slices	Rotation implemented (stereo pairs per second)	Without rotation (stereo pairs per second)
1	103.3	103.3
3	44.4	44.4
5	33.7	33.7
9	20.1	20.1
15	13.2	13.2
21	10.1	10.1
33	6.6	6.6
45	5.0	5.0

Table 2. Frame rates measured as stereo pairs per second for rendering on GPU card with and without rotation implementation.

Appendix D

Real-time stereographic display of volumetric datasets in radiology

Xiao Hui Wang*, Glenn S. Maitz, J. Ken Leader, Walter F. Good
Imaging Research, Dept. of Radiology, University of Pittsburgh, Pittsburgh, PA USA 15213

ABSTRACT

A workstation for testing the efficacy of stereographic displays for applications in radiology has been developed, and is currently being tested on lung CT exams acquired for lung cancer screening. The system exploits pre-staged rendering to achieve real-time dynamic display of slabs, where slab thickness, axial position, rendering method, brightness and contrast are interactively controlled by viewers. Stereo presentation is achieved by use of either frame-swapping images or cross-polarizing images. The system enables viewers to toggle between alternative renderings such as one using distance-weighted ray casting by maximum-intensity-projection, which is optimal for detection of small features in many cases, and ray casting by distance-weighted averaging, for characterizing features once detected. A reporting mechanism is provided which allows viewers to use a stereo cursor to measure and mark the 3D locations of specific features of interest, after which a pop-up dialog box appears for entering findings. The system's impact on performance is being tested on chest CT exams for lung cancer screening. Radiologists' subjective assessments have been solicited for other kinds of 3D exams (e.g., breast MRI) and their responses have been positive. Objective estimates of changes in performance and efficiency, however, must await the conclusion of our study.

Keywords: volumetric display, stereoscopic, lung CT, graphical user interface

1. INTRODUCTION

With rapidly evolving technology in medical imaging, radiology is shifting from 2-dimensional (2D) projective images to 3-dimensional (3D) volumetric datasets. This transition of data acquisition and representation has considerable impact on the medical image interpretation processes employed in traditional practice of radiology. The challenge of 3D volumetric medical data comes from increasing data load and the need for observers to mentally integrate multiple images in order to appreciate 3D structure. A lung CT scan, for example, typically produces about 100 image slices with 2.5-mm collimation reconstruction, and more than 200 image slices when reconstructed at a slice thickness of 1.25-mm. The time and effort to examine a case is roughly proportional to the number of images to be reviewed. Also, with a slice-by-slice viewing method, radiologists need to perform a rather tedious task in which they mentally reconstruct 3D data volumes, and then interpret this mental picture, while navigating through the dataset. Furthermore, in order to maintain a constant x-ray exposure that is independent of the number of slices in exams, the signal/noise ratio of each image will likely be reduced as the number of images generated (i.e., axial resolution) for each case is increased. This can greatly reduce reader's ability to detect features in individual slices.

To more efficiently utilize and present 3D datasets, it is becoming increasingly common for slices to be combined in a manner that can improve signal/noise ratios and, at the same time, give some appreciation of the 3D structure. In most 3D visualization tasks, the appearance of 3D is achieved by applying either surface or volume rendering methods, and projecting the renderings onto 2D displays. Such methods are often unable to depict the spatial relationships of objects without applying dynamic motion or lighting models that unduly affect the diagnostic quality of the dataset. While dynamic displays, such as those incorporating rotation of the volume, have been used to provide a sense of depth, these suffer from impaired performance due to the fact that visual acuity is reduced when viewing objects in motion. In addition, surface rendering methods involve segmenting objects and this can produce unacceptable artifacts due to difficulties in unambiguously identifying surface voxels.

The use of stereographic displays for volumetric data has the potential to overcome many of the limitations of previous display methods. These displays provide simulated views corresponding to what the left and right eyes would see when viewing the data volume and, in doing so, exploit stereopsis, a natural mechanism used by human visual system for depth perception in 3D environments. For observers with normal vision this is a much more natural way to view 3D data.

Stereo workstations meeting the real-time performance requirements for the display of diagnostic medical images are not generally available and this has limited our ability to test the efficacy of stereoscopic displays for clinical applications. For a display to be viable, it must allow observers to dynamically manipulate stereo views in real time, while providing full-resolution stereo renderings. Furthermore, observers should be provided with a variety of appropriate projection methods, because choice of an optimal method depends on the specific clinical task.

We have developed a stereoscopic display workstation designed specifically to meet the requirements for the display of CT and MRI datasets, for the purpose of performing observer performance studies. This workstation provides for the stereo display of sliding thick slabs comprised of multiple CT or MRI slices. Users are able

* xwang@mail.magee.edu; phone 412 641 2561; fax 412 641 2582; www.pitt.edu

to adjust slab thickness, the axial position of slabs, and the rendering method, while renderings appear in real time. Because our immediate goal was to test the efficacy of stereographic methods, we also included functionality to record findings and run observer performance studies. This workstation is currently being used to measure observer performance of stereo displays relative to performance achieved on traditional displays, for CT datasets acquired in a lung cancer-screening project. The remainder of this manuscript provides a description of the design and performance characteristics of the workstation.

2. MATERIALS AND METHODS

2.1 Hardware integration

Off-the-shelf hardware components were integrated to implement a display system that provides for frame-swapped stereoscopic renderings, which users can view through shutter glasses, as well as for superimposed cross-polarized renderings that are viewed through passive cross-polarizing glasses. Frame swapping was included because it has been the primary method for displaying high-quality computer rendered stereo projections in the past, and cross-polarization because recent implementations have demonstrated superior image quality over frame swapping on CRTs. The hardware for our display consisted of a PC computer equipped with a stereographic video card, a programmable keypad, and the two mechanisms for stereo presentation. The computer's performance and stereographic capability have been tested for the visualization of various medical 3D datasets. Though such configurations for stereo display are fairly common, the details are presented below because, in this application, performance depends significantly on the specific configuration.

Computer – A 2.8 GHz AMD Athlon 64 personal computer was configured for stereo rendering and display. The computer has 3 hard disk drives connected via RAID technology to create high-speed disk capacity of 400 GB that can support the large volume of image data required to achieve real-time display, by precalculating renderings as described below. A programmable graphics adapter (nVidia™ Quadro™ FX1100), that provides four display buffers and includes Open Graphics Library (openGL)/DirectX support, was installed for generating the frame swapped display.

Programmable Keypad – User interaction with the display is through the use of a programmable keypad, which is shown in Figure 1. The key-controlled features include, but are not limited to, adjusting window/level settings, changing number of image slices that are used for composing stereo images (i.e., slab thickness), positioning the rendering volume in the axial direction, and toggling visible cues used to mark locations. This was designed to enable radiologists to control the more common display features with one hand while viewing images.

Frame-swapped stereo – In order to achieve stereopsis, left and right images need to be viewed separately by the corresponding eyes, at a frame rate of at least 60 HZ for each eye. We adopted shutter glasses and an emitter, for synchronizing shutter-glasses to the frame refresh signal, from StereoGraphics, Corp. The emitter connects directly to the video card and transmits an infrared light pulse that triggers the glasses to alternate between eyes, and makes it possible to use the shutter glasses without having wires attached. The graphic card was configured to alternately display left and right images, and output a signal through the emitter to synchronize the shutter glasses to the refresh cycle of the monitor. The monitor with refresh rate of 144 HZ is used to give flicker-free stereo images viewed through shutter glasses.

Cross-polarization stereo – A display was provided by Planar Systems (Beaverton, OR) that superimposes polarized images from two digital monitors, by exploiting a special beam-splitter (i.e., StereoMirror™) to combine the images. This has the advantages that contrast and brightness characteristics are somewhat better than on CRT displays, and that both images are displayed at the full frame rate of the monitors, which completely eliminates flicker. This method of stereo presentation uses the dual output of the video adapter but does not require frame swapping.

2.2 Software for display

To test the relative performance of a stereo display, as compared to more traditional display methods, it is necessary to provide for real-time user interaction. Because workstations having sufficient computational power to render stereo projections of CT data in real-time are not generally available, to achieve the desired performance on our workstation we implemented a mechanism to precalculate renderings. The precalculated renderings can be displayed in real-time in response to user interaction. This is not a severe drawback in the CT and MRI applications for which the workstation is intended because these reading tasks involve only a limited set of views (e.g., based on only axial position and slab thickness) which make it possible to transparently provide real-time interaction. The software is comprised of a program for prerendering and displaying projections, and a user interface that enables users to interact with the display.

2.2.1 Routines for prerendering projections

Three rendering methods were implemented. Stereo renderings were by distance-adjusted averaging and distance-adjusted Maximum Intensity Projection (MIP), and for comparison, we also included traditional monoscopic MIP renderings. The monoscopic slice-by-slice display mode is actually a limiting case (i.e., volume thickness = 1 slice) of each of the other modes, so is always available. The details of the rendering calculations have been presented elsewhere [1-3]. Stereo pairs corresponding to all admissible combinations of slab thickness and axial position are precalculated, organized into a 2-dimensional linked list that allows projections to be accessed by thickness and axial position, and stored on the display's hard disk.

For our ongoing lung nodule study, renderings are calculated for slab thicknesses ranging from 1 to 19 slices. For each thickness, slabs are positioned along the axis at intervals equal to 20% of the slab thickness. This process requires approximately 1-GB of disk storage per case, but the method is able to achieve update rates of greater than 5 updates per second in response to changes in axial position or slice thickness.

The actual display of stereo pairs is managed through openGL and uses the two double buffers on the nVidia™ display adapter for the left- and right-eye views. Frame swapping is automatic once stereo display is enabled through openGL.

2.2.2 User interface

The user interface software is written as a Windows application, implemented with Win32 and MFC (Microsoft Foundation Classes). The Windows API provides a Windows framework and functional utilities for the user interface. Stereographic display is enabled by using the nVidia™ version of openGL functions to manipulate display functions on the graphic processing unit.

The user interface enables viewers to navigate in real-time the entire volumetric dataset, to change slab thickness, and to adjust the brightness/contrast of displayed images, but does not allow for oblique views to be calculated in real time. All user input is by means of a mouse and the keypad. The mouse is used for marking locations on the image and for responding to questions on scoring forms during performance studies. All image manipulations are controlled through the keypad.

For marking locations in 3D displayed volumes, a stereo cursor has been implemented. This cursor, which is controlled by the mouse, is projected by the same perspective transform as that applied to the data. Thus, the size of the cursor is properly transformed to reflect the size perceived at a distance based on the spatial location of the cursor in the volume, and the horizontal disparity between the left- and right-views are appropriately adjusted to reflect depth. The cursor also incorporates a depth-sensitive scale for measuring the sizes of features in the 3D volume.

A window/level (contrast/brightness) mechanism was built into the display workstation with initial default settings for different tissue types, such as soft tissue, bone, and lung tissue. Real-time adjustment of window/level is enabled through encoded keys on the keypad.

As we previously found in a related study, for certain kinds of features MIP rendering is best for detection because of its high contrast, while rendering by averaging is better for characterization of features once they have been detected because it gives a more accurate representation of local geometry [2-3]. In the display software, we have included a button for toggling between these two rendering methods.

2.3 Software for observer performance study

The intended application of the workstation is to perform observer performance studies to test the efficacy of stereo display versus monoscopic displays for certain clinical tasks involving CT and MRI datasets. Software was written to implement functions required for this kind of study. This involves tasks such as management of cases and display modes, case randomization, electronic scoring of detected nodules, and data archive. All software was written in the C++ language and rigorously tested.

2.3.1 Implementation of case randomization

In observer performance studies, case randomization is a procedure to assure that a reader's response is not biased by the order of case presentation. For a given reader and reading mode, the randomization process begins by selecting all cases that have not previously been seen by the reader in the specified mode, and eliminates all cases that the reader has seen in any mode within a pre-designated time interval, to eliminate the possibility that the reader will remember the case from a previous reading session. The remaining list of cases is randomized for presentation to the reader.

2.3.2 Recording of findings

An electronic reporting mechanism is provided which allows viewers to use a stereo cursor to measure and mark the 3D locations of specific features of interest, and then record findings. When a reader clicks on an image feature of interest, a scoring form with study questions will appear on the screen. The questions, which are normally in the form of check boxes, radio buttons and sliders, can be tailored to the specific requirements of individual studies. All answers, as well as the location of the identified feature, are saved after the reader finishes scoring a case. A display of cues marking the locations of detected features is provided and can be toggled on and off at reader's will.

2.3.3 Routines for monitoring reading patterns

Routines have been incorporated into the study software for keeping track of time and changes in viewing parameters made by radiologists when viewing a case. The display program periodically (i.e., every 5 milliseconds) records reading status, including location of displayed slab, slab thickness, and rendering method in the case of stereo display modes. This data will be a part of the evaluation criteria for display studies and is expected to become a valuable reference for understanding psychophysical attributes in different display designs.

3. RESULTS

In order to evaluate clinical efficiency of our stereo display workstation, we have been conducting an observer performance study of stereo display of chest CT for lung nodule screening. In that study, radiologists are asked to identify possible nodules, mark their locations, and then report findings in a scoring form, while viewing the dataset in one of three display modes. While this study is in an early stage, and its results are not the subject of this manuscript, we will use it as a demonstration of how the workstation performs in practice. A typical screen display for this study is shown in Figure 2 and the corresponding scoring form is shown in Figure 3.

In general, the functionality described in the methods above works as intended. The system has achieved real-time dynamic display of slabs, where slab thickness, axial position, rendering method, brightness and contrast are interactively controlled by viewers. Viewers were unaware that projections had been precalculated.

Using our prerendering strategy, we are able to achieve a stereo update rates in excess of 5 updates/sec. Radiologists were comfortable performing image interpretation at this rate and seldom report unusual fatigue or other adverse effects from use of the display, however some participants have not been fond of the requirement to wear shutter-glasses or cross-polarized glasses, both of which can create problems when viewing surroundings.

Both the frame-swapped and cross-polarized displays exhibit some amount of ghosting (i.e., crosstalk between the left- and right-eye views). For the frame-swapped display, this is due to the fact that the persistence of the phosphor is long enough that the display on one view may not have sufficient time to completely decay before the alternate view is displayed. Cross-polarization suffers from the same kind of artifact, but it is caused by a different process. If the polarizations of the displayed images are not orthogonal, or the polarization filters in the glasses are not orthogonal, or the glasses

are not properly aligned with the polarizations of the images, then it may not be possible for a filter in the glasses to completely block the opposing view. In this case, ghosting is dependent on the degree of head tilt relative to the display. Overall, ghosting is less of a factor on cross-polarized displays than on frame swapped displays using current CRT technology.

4. DISCUSSION

Despite the potential of stereo display methods to have a significant impact on radiology, little information exists on the relative performance that might be achieved on stereo displays versus that on monoscopic displays. The main reason for this is that workstations having sufficient performance for stereo display, including real-time manipulation of views, have not been readily available. Current medical image workstations for viewing volumetric datasets acquired by MRI or CT need functionality that goes well beyond the simple display of images in the traditional slice-by-slice mode [4]. The workstation discussed herein was designed specifically for performing the kinds of studies required to establish the viability of stereoscopic display in radiological applications.

The problem of ghosting, which is evident in both of the stereo display technologies, is an important issue when stereo is provided on radiographic displays. The problem could be largely overcome on frame-swapped CRT displays by adopting a shorter-persistence phosphor, but such displays are not commercially available. For cross-polarized displays, there will be considerable benefit in training observers to recognize ghosting and then to eliminate it by properly orienting their eyes relative to the display.

A fully functional stereo display for radiological applications will require that readers be able to dynamically choose arbitrary views, and have them rendered in real time. This involves real-time computation in addition to real-time display. The necessary computational power, on reasonable personal workstations, is just becoming available. Recent commodity programmable graphics processor units, which are designed for rendering computer games, are readily available and can be programmed to perform the kinds of real-time calculations that are needed in stereo workstations. This work is ongoing, by us and by others, and will eventually replace the precalculation mechanism described above and, in doing so, will allow arbitrary views and renderings.

New display technologies are rapidly being introduced that do not require the use auxiliary devices to view stereo displays. At the present time, the image quality and viewing requirements of these autostereoscopic displays are not adequate for radiological applications, but their performance is continuing to improve. An alternative for the future are true volumetric display, which do not attempt to simulate stereoscopy by rendering left- and right-eye views, but rather display the data over a 3D volume in space. While displays of this type exist in laboratories and have been applied to a few specialized applications, they are still at a very early stage of development and are not yet ready for application in radiology.

As experienced by many radiologists, stereo displays depict complex anatomy better than monoscopic 3D displays. There is also an indication that they are more efficient for viewing CT and MRI datasets than traditional slice-by-slice displays. Nevertheless, the actual clinical utility of our stereoscopic display workstation still needs to be established through extensive observer performance studies.

ACKNOWLEDGEMENTS

This work is sponsored in part by grant CA80836 from the National Cancer Institute, National Institutes of Health, and also by the US Army Medical Research Acquisition Center, 820 Chandler Street, Fort Detrick, MD 21702-5014 under Contract DAMD17-02-1-0549 and contract PR043488. The content of the contained information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred. The authors also thank Planar Systems, Inc., of Beaverton, OR, for loaning us the cross-polarization display used in this project.

REFERENCES

1. X. H. Wang, W. F. Good, C.R. Fuhrman, et al., "Stereo Display for Chest CT", Proc SPIE, **5291**:17-24, 2004.
2. X. H. Wang, W. F. Good, C. R. Fuhrman, et al., "Projection Models for Stereo Display of Chest CT", Proc SPIE, **5367**:676-686. 2004.
3. X. H. Wang, W. F. Good, C. R. Fuhrman, J. H. Sumkin, C. A. Britton, S. K. Golls, "Stereo CT compositing methods for lung nodule detection and characterization", Academic Radiology, 2005, (in press).
4. A. Moise, M.S. Atkins, "Designing better radiology workstations: impact of two user interfaces on interpretation errors and user satisfaction", J Digit Imaging. **18**:109-115, 2005.

FIGURES

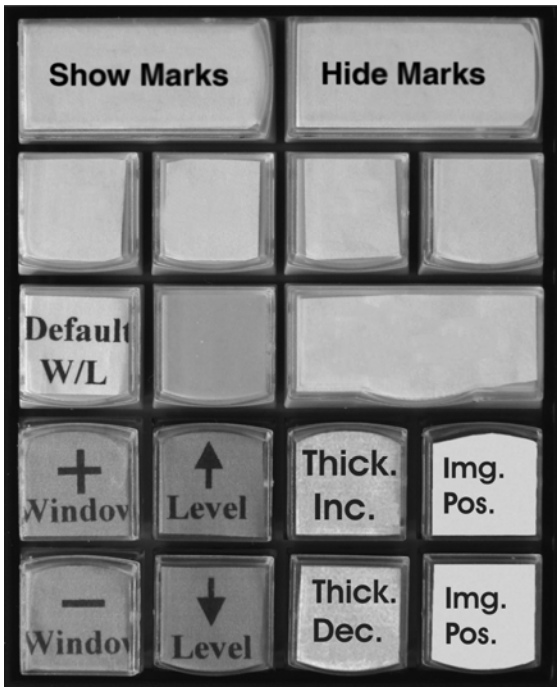


Figure 1: Programmable keypad with which user can control the main display features.

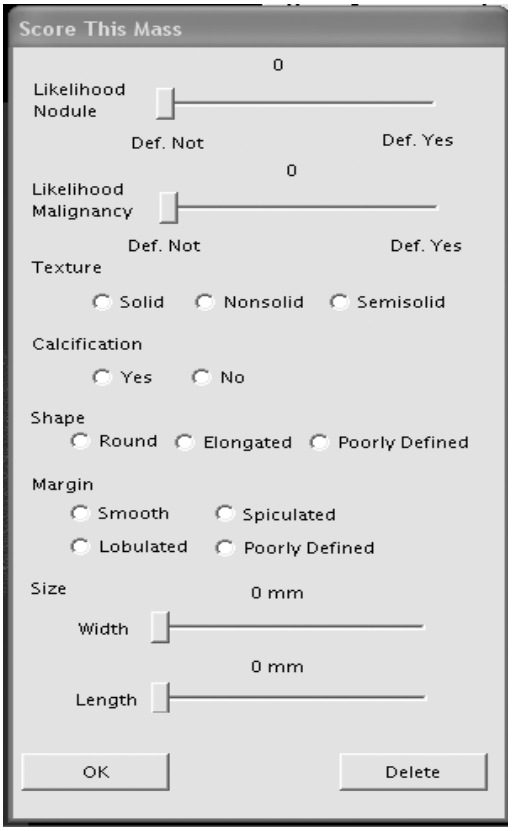


Figure 3: Scoring form used to enter data in current lung nodule study.

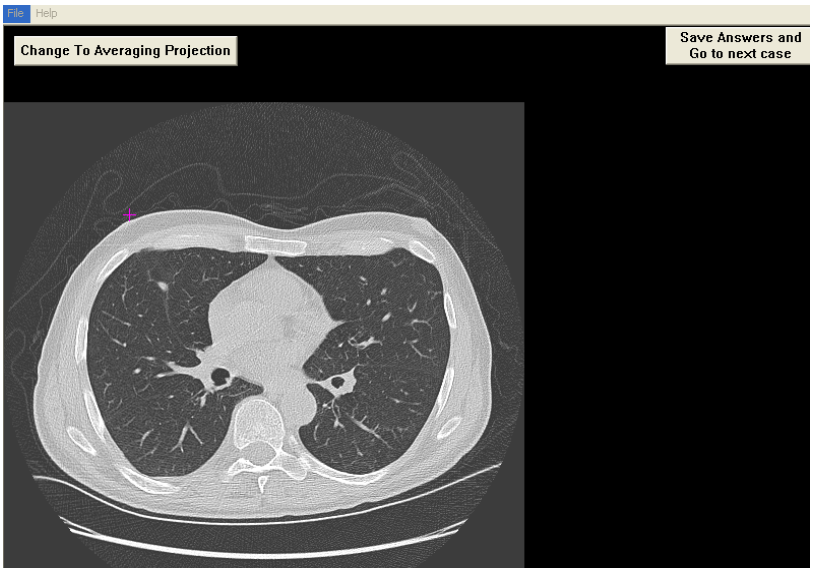


Figure 2: Window-based user interface.

Appendix E

Stereo display of CT images for lung cancer screening: a pilot study

Xiao Hui Wang, Janet E. Durick, David L. Herbert, Amy Lu, Saraswathi K. Golla, Dilip D. Shinde, Samaia Piracha, Kristin Foley, Carl R. Fuhrman, Betty E. Shindel, J. Ken Leader, Walter F. Good

Department of Radiology, University of Pittsburgh

ABSTRACT

To improve radiologist's performance in lesion detection and diagnosis on 3D medical image dataset, we have conducted a pilot study to test viability and efficiency of the stereo display for lung nodule detection and classification. Using our previously developed stereo compositing methods, stereo image pairs were prestaged and precalculated from CT slices for real-time interactive display. Three display modes (i.e., stereoscopic 3D, orthogonal MIP and slice-by-slice) were compared for lung nodule detection and total of eight radiologists have participated this pilot study to interpret the images. The performance of lung nodule detection was analyzed and compared between the modes using FROC analysis. Subjective assessment indicates that stereo display was well accepted by the radiologists, despite some uncertainty of beneficial results due to the novelty of the display. The FROC analysis indicates a trend that, among the three display modes, stereo display resulted in the best performance of nodule detection followed by slice-based display, although no statistically significant difference was shown between the three modes. The stereo display of a stack of thin CT slices has the potential to clarify three-dimensional structures, while avoiding ambiguities due to tissue superposition. Few studies, however, have addressed actual utility of stereo display for medical diagnosis. Our preliminary results suggest a potential role of stereo display for improving radiologists' performance in medical detection and diagnosis, and also indicate some factors likely affect the performance with new display, such as novelty of the display, training effect from projected radiography interpretation and confidence with the new technology.

Keywords: Stereo display, Lung CT, Visualization, Validation, FROC

1. Introduction

Lung cancer is a leading cause of cancer death in both men and women in the United States. In several large lung cancer screening trials, low-dose helical computer tomography (CT) has proven to be an effective tool for lung cancer screening with superior sensitivity (80 ~ 90%) for early detection, compared to other methods (e.g., chest radiographs with about 23% sensitivity and sputum cytology with 10 ~ 20% sensitivity) [1-3]. Despite the advantages of using CT as a screening tool for lung cancer, current methods for displaying and interpreting chest CT data are inefficient and inadequate.

At present, the two most common viewing methods employed by radiologists for interpreting these studies involve either reading individual images in a sequential slice-by-slice mode, or viewing thicker slabs comprised of multiple sequential slices projected onto a 2D display. The slice-by-slice method makes it necessary for radiologists to reconstruct mentally 3D information represented in sequential 2D slices in order to differentiate between nodules and linear structures, such as blood vessels, passing through the slices. In addition, the signal-to-noise ratio in single slices may be too low for subtle lesions to be reliably detected.

We have proposed the use of stereoscopic 3D displays for reading lung CT images in hope that this will increase both efficiency and detection performance beyond what can be achieved with other display methods [4,5]. Such displays, which have the potential to provide more natural representations and volume-based views of structures for viewers having normal binocular vision, have been studied in the past for the display of medical images under various circumstances but, due primarily to technical limitations (e.g., computational power, display technology), these methods have not been widely adopted.

Lung CT images are well suited for stereo viewing because, by making air transparent, the sparsely distributed lung tissues of interest (e.g., vessels, airways and nodules) can be easily visualized. The lung geometry, at the inspiration level maintained during screening exams, makes it especially easy to create unambiguous projections along the axial direction. Stereoscopic displays have the potential to increase efficiency and signal-to-noise ratios by enabling the display of thicker tissue volumes.

It is known that certain kinds of objects can be detected in a stereo 3D display of data, which cannot be detected when the data is viewed in a slice-by-slice manner. Stereo projection can improve the visibility of objects by enhancing features that are correlated between slices, while reducing noise, in a manner analogous to the signal-to-noise improvements obtained by averaging slices or traditional MIPs – but stereo projection does not introduce tissue superposition ambiguities that would be caused by these methods. In addition, there are situations where object detection depends on stereopsis (e.g., random dot stereograms). Consequently, it is often possible to detect objects on stereo 3D displays that cannot be seen when the same data is viewed in a slice-by-slice manner [6,7].

The objective of this project is to develop and evaluate methods for displaying, on stereoscopic displays, lung cancer screening data acquired by helical CT. In particular, we hope to determine whether these displays have the potential to improve the efficiency and accuracy of radiologists in reading these studies.

Findings from the limited number of relevant studies indicate that complex structures were better observed with stereo display and, moreover, radiologists' detection tasks were performed more efficiently using a stereo display mode [8,9]. Also, there is evidence, though preliminary, that radiologists prefer to view 3D datasets on stereo displays [10].

The current trend toward acquisition of large 3D datasets combined with improvements in the relevant technologies for stereographic display has led us to reconsider the potential role of these displays for radiology. As part of this effort, a display system was developed and then tested, in a small pilot study.

2. Methods

A workstation was developed that was capable of displaying CT datasets stereoscopically as well as by MIP and in a traditional slice-by-slice mode. This workstation was then evaluated in a small pilot study that was designed to provide the essential information required to design a more comprehensive study.

2.1. Workstation – The physical display consisted of a CRT viewed through shutterglasses, which allowed each eye to see a different view of the data. The workstation is depicted in figure 1. A keypad was provided, as shown in figure 2, by which readers were able to adjust window, level and axial position of the display within the dataset. It was also possible in the Stereo and MIP display modes to adjust slab thickness by use of the keypad. All display controls operated at real-time rates.

2.2. Pilot study – The pilot study was organized as a retrospective study of 30 cases, half containing masses, which were interpreted in each of three display modes by 6 board certified radiologists, who had extensive experience in reading chest CT.

Protocols for managing this sort of performance study have been employed in our facility over the past 15 years and have been widely reported [11]. In this study, three display modes were tested: 1) slice-by-slice; 2) thick-slice MIP; and, 3) 3D stereoscopic. The slice-by-slice mode was included because of its traditional stature, while thick-slice MIP was included because it is currently available on some commercial systems. The basic task was to detect and classify nodules. Each case was interpreted, in each mode, by six board certified radiologists, who reported on the likelihood of the existence of a nodule, the likelihood that a suspected nodule is malignant, lesion size and their subjective evaluation of the display. The readers marked the location of suspected lesions in the 3D dataset by using a 3D cursor with a 1 cm built-in scale for measuring size. Once a suspected lesion had been marked, a computerized scoring form popped up with the questions that were to be answered. The order of modes, as well as the order of cases was randomized and counterbalanced with the constraint that the same case could not be read by the same reader without a delay of at least two weeks. During a given reading session, all cases were displayed in the same mode. In both the 3D stereoscopic and thick-slice MIP modes, readers were able to adjust the thickness of the displayed lung volume. Stereo mode cases were initially present with a rendering that is considered to be optimal for nodule detection, but the reader had the capability of changing to a rendering that was more appropriate for characterizing nodules.

By design, in both the stereo and thick-slice MIP modes, readers had the option to view images at a thickness of one, which is equivalent to viewing the data in the slice-by-slice mode. This was allowed in order to make the study more relevant to the clinical environment, where the radiologist will always need to have the original slice data available. To insure that readers used the thicker modes when they were available, the system was designed to not allow single slices to be viewed until all data has first been displayed as a thicker volume. While this constraint is a slight departure from what would be desirable in the clinical environment, it seems to us to be appropriate, in this situation, because of differences in readers' experience (and comfort) between the slice-by-slice mode and the two thick slice modes. In any case, we recorded axial position during the study and this was monitored as part of our quality assurance program.

The data collection strategy was designed for analysis by FROC methods. In addition to the FROC type of questions, we asked readers to report their subjective assessment of the appropriateness of the current display mode, for each case. The time required to read each case was recorded in order to estimate relative efficiencies of the various modes. Axial positions within image datasets, and image thickness were also automatically recorded during reading sessions.

All readers who participate in the project were Board certified radiologists, with a minimum of three years' experience in the interpretation of chest CT exams. They were not aware of the aims of the study in which they participated. Participating radiologists received an "Instructions for Observers" form for review, and the definition of abnormalities was discussed with each reader.

2.3. Case selection and verification – Helical CT images used for developing and evaluating the display were obtained from subjects who have previously been scanned as part of the Pittsburgh Lung Cancer Screening Study – an ongoing lung cancer screening trial of subjects that are considered to be at high risk due to age and smoking history. Cases were acquired on a helical CT scanner (LightSpeed Plus, GE Medical Systems, Milwaukee, WI) using X-ray tube current of 40-mA, voltage of 140-kVp and 0.5-mm pitch. All images were acquired in the axial plane and reconstructed to a thickness of 2.5 mm/slice with GE standard convolution software for lung tissue. The pixel size in each slice is 0.75-mm × 0.75-mm.

For the pilot study we collected 35 cases (i.e., 5-training, 30-pilot study) half containing at least one nodule. Half of the nodule cases were malignant. All the nodules used in this paper were identified, verified, marked and characterized by three experienced radiologists. The verification process was repeated at least one time to ensure agreement on the nodule identification and characterization. Positive cases had been verified by positive

biopsy, and negative cases by disease free follow-up, as has been discussed in reference to our previous projects [12].

2.4. Rendering methods – Three display modes were implemented on the workstation: 1) slice-by-slice; 2) thick-slice MIP; and, 3) 3D stereoscopic.

2.4.1. Slice-by-slice display **–** The slice-by-slice mode was include because it is the most basic mode and is universally available of CT displays. It required only minimal additional software to implement because, by design, this mode is actually a limiting case (i.e., volume thickness = one slice) of both the MIP and stereo modes.

2.4.2. Stereographic rendering **–** For CT data to be displayed stereoscopically, a 3D dataset must be projected onto two views, corresponding to observers' left and right eyes. This involves using a geometric perspective transformation which projects rays through the 3D volume onto individual pixels. These projection methods have been studied extensively, mostly within the computer graphics literature [67]. Stereo image pairs were generated by stereo projection algorithms that incorporated transparency-contrast models optimized for this application. Specifically, the rendering models were designed so as to make air transparent, and maintain high contrast for nodule detection, and geometric fidelity for nodule characterization.

The visual characteristics of the projected image depend on the specific raycasting method employed. Previously we have shown that for stereo rendering, because of the tradeoff between contrast for detection and geometric fidelity for characterization, raycasting based on MIP performs better for detection tasks while raycasting based on averaging is preferable for assessing features once they have been detected [5]. For this project we provided both of these raycasting methods for stereo projection and observers were able to switch between them at will. Because nodules must be detected before they can be characterized, we believe that the use of two separate stereoscopic display modes is both desirable and feasible. Once a nodule had been detected in a stereo MIP rendering, we expected that the reader would adjust the thickness of the displayed volume to include only those slices that contain the nodule, and redisplay this volume using the raycasting by averaging method in order that the nodule can be more accurately characterized.

In the conventional geometric perspective transformation that was used to compose left-eye and right-eye image pairs, we adopted an interpupilary distance of 6.5-cm, a viewing distance (the distance between eyes and the screen) of 45-cm and a display area of 25-cm × 25-cm. The perspective transformation was symmetrical, based on the assumption that a viewer is centered in front of the screen.

The voxel shape in the CT images used in this study was nonisotropic in that images had been reconstructed to a larger thickness in z direction (i.e., slice thickness of 2.5-mm) than in x and y directions (pixel dimension: 0.75 × 0.75-mm). To approximate isotropic voxels, three slices were created by trilinear interpolation between each pair of adjacent CT slices. Consecutive slices, including CT and interpolated slices, within a given volume were used for generating a stereo pair.

Both raycasting processes used a light emission / transmission / occlusion model that assumed that each voxel emits light in proportion to its voxel value when the CT images are displayed at a normal window and level for nodule detection, and that uses distance information (distance weighing factors) to determine the amount of this emitted light that reaches the projection plane. Specifically, it was assumed that each slice has a fixed optical density that reduces the brightness of slices lying behind it. The total of all distance-weights was equal to one. The ratio of the weights between the last slice (the slice with the largest distance from screen) and the first slice (the slice at screen level) controls level of transparency for a given volume. We have studied a range of these ratios for lung CT images, and empirically set the ratio to 0.5 in order to achieve a balance between the use of brightness weighting as a depth cue and the visibility of the back slice. The final value for a voxel is the sum of distance-weighted pixel values in a perspective transformation ray. The detailed calculations were described in references [4] and [5].

2.4.3. Monoscopic MIP rendering **–** For the comparisons in the study, it was also necessary to generate traditional monoscopic MIP images. To be consistent with commercially available CT displays, these were generated using a standard orthogonal projection. In this mode, readers were able to change slice thickness as well as axial position by using the keypad.

3. Results

3.1. Detection performance – Table 2 presents the detection results that were obtained by FROC analysis. As can be seen from the table, while there are apparent differences between modes, none of these differences reached the level of statistical significance. This was not unexpected because the small size of the pilot study did not provide sufficient power to detect small changes in detection performance.

Table 1. Figure-Of-Merits from the FROC curves for lung nodule detection performance.

Stereo	Orthogonal MIP	Slice by slice
0.57	0.52	0.56

3.2. Reading time – Table 3 indicates the case-based interpretation time averaged from all participating radiologists for each display mode. There was less average time required for the interpretation performed on stereo display mode than on either MIP or slice-by-slice mode.

Table 2. Average interpretation time.

	Stereo	Orthogonal MIP	Slice by slice
Time (minute/case)	3.5	3.7	4.5

3.3. Viewing strategy – We found that, while readers spent some time viewing the data at a thickness of one, most of their time was spent viewing thicker slices.

4. Discussion

4.1. Stereo by MIP projection – Because of its ability to solve the contrast problem in many situations, MIP has been widely adopted for the monoscopic display of thicker slabs. However, it is not *a priori* clear that MIP is suitable for stereo compositing because of its inability to preserve the local geometric features and texture of objects and, to a large extent, it is this local information that is required for stereopsis. To achieve stereopsis from two views, a viewer needs to detect corresponding features in each view and to determine the relative geometric disparity between those features. But with MIP projections, features that appear in one view (i.e., are a maximum along projections onto the view) may not be represented in a different view if they are not of maximum intensity along a projection for that view, and such features may provide misleading geometric disparity information. This causes small bright objects to have a specular appearance. However, MIP does preserve the presence, but not the exact geometry, of sparsely distributed clusters of bright voxels when they are viewed against a darker background, which is essentially the situation that usually occurs when a nodule is displayed in a thick slab from axial CT slices of the lung. For these reasons, images produced by MIP are, in principle, in conflict with plausible psychophysical transparency/brightness and geometric vision models, but still were generally able to provide high-contrast for detection in the context of this project.

4.2. Ergonomics of stereo display – There are several specific concerns, with respect to stereo display, that could only partially be addressed in this project.

If a certain level of display performance is not achieved, in terms of display refresh rate and image quality, prolonged viewing of stereo images on a CRT can result in excessive fatigue, eyestrain and symptoms similar to simulator sickness. In this project, this was largely overcome by employing a high-quality, high-resolution CRT that operated at a refresh rate of 125 Hz. Nevertheless, when CRTs are used in conjunction with shutter glasses, the persistence of the phosphor becomes an issue at these high refresh rates. Specifically, if the phosphor excitation decays too slowly, then each eye will see a ghost of the image intended for the opposite eye. This was an issue in this project though we believe the amount of ghosting was too small to affect detection performance. In the future we will replace this technology with a display based on viewing cross-polarized images superimposed from two liquid-crystal displays, and this should alleviate flicker and ghosting problems.

Stereoscopic displays have traditionally required viewers to wear special glasses or possibly head mounted displays. The need to wear shutterglasses in this project was mentioned as a drawback by 1 of our readers but did not seem to cause problems in our reading environment where they were not switching between display modes in a given session. Various kinds of autostereoscopic displays that can be viewed without the user wearing any special device [14] have recently been developed but have not achieved the level of performance that would make them viable for radiographic display at this time.

A more intrinsic problem is that when stereo images are displayed on a screen, the eyes must focus (accommodate) on the screen, but converge on points at different distances, based on the geometry assumed in the stereo projection. This causes an inconsistency in the learned response that associates accommodation with convergence, and can result in eyestrain. Our experience is that readers slowly adapt to this, but readers in the pilot study did often express a sense of fatigue.

4.3. Learning curve – The relative novelty of stereo display makes it necessary for readers to be trained before a meaningful comparison between display modes can be achieved. We found that readers initially did a significant amount of experimentation to find an optimal slab thickness and viewing strategy before they settled into a consistent pattern.

5. Conclusion

While the results of the pilot study were limited by the small size of the study, they were consistent with our hypothesis that there are certain advantages to stereo display. Specifically, the trend was that the nodule detection on stereo display outperformed the one on slice-by-slice mode or on conventional MIP display mode (table 1), though this was not statistically significant; and in terms of interpretation time, stereo display was the most efficient among the three display modes for the lung nodule detection (table 2).

Based on the limited results of this pilot project and our continuing belief that there are many advantages to stereographic displays, and that the limitations can be overcome, we remain optimistic that these displays can improve readers' performance and efficiency for many tasks involving 3D datasets.

ACKNOWLEDGEMENTS

This work is sponsored in part by grant CA80836 from the National Cancer Institute, National Institutes of Health, by Planar Systems, Beaverton, OR, and also by the US Army Medical Research Acquisition Center, 820 Chandler Street, Fort Detrick, MD 21702-5014 under Contract DAMD17-02-1-0549 and contract PR043488. The content of the contained information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

REFERENCES

1. Kubik A, Polak J. "Lung cancer detection: results of a randomized prospective study in Czechoslovakia", Cancer, **57**, 2427-2437, 1986.
2. Fontana RS, Sanderson DR, Taylor WF, Woolner LB, Miller WE, Muhm JR, Uhlenhopp MA. "Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Mayo Clinic study", Am Rev Resp Dis., **130**, 561-565, 1984.
3. Fontana RS, Sanderson DR, Woolner LB, et al. "Screening for lung cancer: a critique of the Mayo lung project", Cancer, **67**, 1155-1164, 1991.
4. Wang XH, Good WF, Fuhrman CR, et al. "Stereo Display for Chest CT", Proc SPIE, *5291*, 17-24, 2004.
5. Wang XH, Good WF, Fuhrman CR, et al. "Projection Models for Stereo Display of Chest CT", Proc SPIE, *5367*, 676-686, 2004.
6. Smith PA, Marshall FF, Urban BA, Heath DG, Fishman EK. "Three-dimensional CT stereoscopic visualization of renal

masses: impact on diagnosis and patient management", AJR, **169**, 1331-1334, 1997.
7. Brown DG, Riederer SJ. "Contrast-to-noise ratios in maximum intensity projection images", Magn. Reson. Med., **23**, 130-137, 1992.
8. Hsu J, Chelberg DM, Babbs CF, Pizlo Z, Delp EJ. "Pre-clinical ROC studies of digital stereomammography", IEEE Trans. Med. Imaging, **14**, 318-327, 1995.
9. Sakakura A, Yamamoto Y, Uesugi Y, Nakai K, Hayashi I, Makimoto K, Takenaka H Narabayashi, I. "Stereoscopic display of a three-dimensional image of the larynx using high-speed helical scanning", ORL J Otorhinolaryngol Relat Spec., **62**, 290-295, 2000.
10. Calhoun PS, Kuszyk BS, Heath DG, Carley JC, Fishman EK. "Three-dimensional volume rendering of spiral CT data: theory and method", Radiographics, **19**, 745-764, 1999.
11. Thaete LF, Fuhrman CR, Oliver JH, Britton CA, Campbell WL, Feist JH, Straub WH, Davis PL, Plunkett MB. "Digital radiography and conventional imaging of the chest: A comparison of observer performance", AJR, **162**, 575-581, 1994.
12. Anderson CM, Saloner D, Tsuruda JS, Shapeero LG, Lee RE. "Artifacts in maximun-intensity-projection display of MR angiograms", AJR Am J Roentgenol., **154**, 623-629, 1990.
13. Mortenson, ME. *Computer Graphics Handbook: Geometry and Mathematics*, Industrial Press, New York, 1990
14. Sawaki A, Shimamoto K, Hattori T, Ikeda M, Ishiguchi T, Ishigaki T, Sakuma S. "Three-dimensional image display without special eyeglasses: observation of magnetic resonance angiography using the stereoscopic liquid crystal display", J. Digital Imaging, **14**, 111-116, 2001.

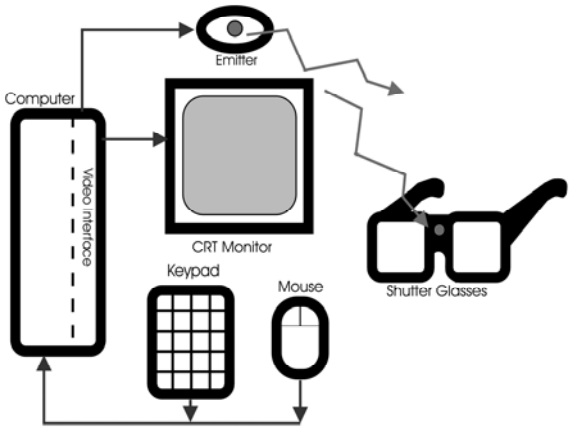


Figure 1: Scematic diagram of stereographic workstation

Figure 2: Keypad used for control of stereographic display

Appendix F

Photometric Correction of Stereographic Image Pairs

Xiao Hui Wang, Walter F. Good; University of Pittsburgh; Pittsburgh, PA, USA

Abstract

Differences in Photometric characteristics between images acquired during stereographic imaging may significantly reduce the effectiveness of their subsequent display or analysis. While uniform global differences can easily be corrected by applying traditional histogram matching techniques, these methods are not capable of dealing with differences that are object or distance dependent. We have developed a procedure to adjust locally, visual characteristics of one image in a stereo pair to match the alternate image. Objects, and their boundaries, are segmented in both images by detecting edges and depth discontinuities, and these features are used to partition the images into connected components. Where possible, stereo correspondences between components in each image are identified and used as the basis for local color correction. The fully automatic procedure is able to remove visible differences in most cases, but further development remains before the system will be sufficiently robust.

Introduction

Typically, stereographic image pairs are acquired on film or digitally, by methods such as synchronized exposures using multiple lenses and detectors; as a single exposure employing some form of beam splitter; or as a sequence of exposures by a single acquisition device which is moved appropriately between exposures. Whatever the acquisition paradigm, errors of either a geometric or photometric nature may occur. These errors can be a result of many factors such as different acquisition times, inadvertent camera motion between acquisitions, film-processing inconsistencies, differences between digital detectors, differences between lenses, and inaccuracies in the relative orientations of lenses. For stereopsis to be easily achieved when viewing stereoscopic image pairs, the geometric relationships between corresponding points must conform to the requirements of epipolar geometry and other visual differences must generally be small.

This manuscript is primarily concerned with correcting for inconsistencies in the photometric characteristics between images. More specifically, it is stipulated herein that for a given stereo image pair, their geometry is correct and one image is considered to be photometrically correct, while the alternate image is inconsistent with the first.

Prior attempts to adjust images in a pair of stereo exposures have generally relied on global methods such as various forms of white balancing or histogram equalization. These methods can be automated in a manner that provides reliable results in most cases, and are used routinely within our laboratory for postprocessing stereo image pairs. While these methods are able to correct for many types of systematic color shifts, they are inadequate for cases where differences are depth dependent or where surface reflectivity is such that an object's appearance changes rapidly or discontinuously with changing angle of view.

To improve our automated postprocessing procedures, we have been investigating object-based methods that attempt to identify corresponding foreground objects and adjust their photometric characteristics if there are significant disparities. This task is closely related to the problem of determining depth from stereo correspondence, as both involve matching objects between images in a stereo pair - which has proven to be very difficult. In fact, these tasks do not have a theoretical solution in general because it is possible to contrive stereo image pairs having no corresponding regions. In general, stereo image pairs will have some regions which correspond and some that do not, and in this situation there may be considerable ambiguity in how regions are to be identified together. Note that this ambiguity occurs both in computerized analysis as well as when a scene is being viewed by human observers, and is the source of many psychophysical depth illusions. The ultimate goal of this project is to use the information in a stereo pair to adjust the images to make it easier for viewers to achieve stereopsis.

Color Space Disclaimer

Within this manuscript we have deemphasized issues related to the exact color space under consideration and whether the color space has a true metric in the mathematical sense. While this allows us to sidestep many difficult technicalities, the main reason this is desirable is that many of our images are acquired with radiation sources other than light (e.g., x-rays) and assignment of a pseudocoloring is somewhat arbitrary. When these images are displayed they are subject to the constraints of color spaces and perceptual metrics, but the algorithms developed herein are applied to the originally acquired data and issues related to color spaces are not relevant at that point.

Methods

For the purposes of this manuscript, it is assumed that in the image pairs being processed the image planes are coplanar and that scan lines in each image are parallel to their common baseline. These conditions are sometimes specified by saying that the images have been rectified [1]. We also assume that the images

were acquired or artificially generated at projection angles that are reasonably representative of human vision, and that the images are of normal kinds of scenes that have not been specifically contrived to defeat computer algorithms.

Object Segmentation

Each image in a pair is first segmented into coherent regions over which chromanence and brightness vary continuously. The process begins by identifying boundaries with an edge detector based on Cranny's algorithm [2]. Images are then partitioned into connected components over which the continuity constraints are enforced. An attempt is then made to establish stereo correspondence between images by independently matching epipolar pairs of scanlines to derive disparity and occlusion information. Our goal is to identify discontinuities in depth which indicate boundaries of objects. These methods have been proposed in computer vision literature by many [3-5] but the specific algorithm we employ, which was first proposed by Birchfield and Tomasi [6], is cvFindStereoCorrespondence(),contained in the openCV library. In this process, because we are primarily interested in larger forground regions, non-uniform regions and very small objects are either suppressed or combined into larger regions, while larger objects are retained. Not all regions can be unambiguously matched between images by this process. Unless the stereo correspondence can be determined with a high degree of confidence, the program does not attempt to correct the region. For each pair of corresponding connected components, a linear correction function that minimizes the sum-of-squares difference is determined.

We have developed software to implement these methods. The algorithms have a number of thresholds and parameters that need to be specified but otherwise operate automatically.

Luminance-Based Figures-of-Merit of Stereo Correspondence

As part of this work we have been investigating measures that are related to the amount of stereoscopic information contained within stereo image pairs. For a given pair of images, the central question is how closely one can come to generating the images in some sense, from projections of a geometrically viable 3D scene. In general, for scenes consisting of opaque surfaces, this does not have an analytical solution, and the effort required to solve it computationally is prohibitive. Thus, an optimal solution is not known and it is unlikely that there is a single measure that captures all aspects of stereo correspondence. Nevertheless we have devised two luminance-based figures-of-merit to measure different aspects of stereo correspondence. Both of these are most useful for evaluating small changes in images that are known to be related by stereo projection, and both behave somewhat unpredictably for large changes or for random images.

The first method produces a 2D scatter plot of luminance values from corresponding pixel pairs taken from the two views. The scatter plot is then analyzed for its degree of clustering, central tendency, and for certain aspects of its symmetry. This method does not require normalization but can be defeated when it is applied to image pairs that are not related by stereo projection. An example of a scatter plot generated by the method is shown in figure 1. This particular plot was generated for the left-eye and corrected right-eye images in figure 3. Note that the strong preponderance of points along the diagonal indicates that the images are very similar, while the deviation from linearity of the distribution is attributed to stereo disparity.

The second method begins by identifying the closest object appearing in both views. The two images are then aligned horizontally so that this object is in registration between the images and only the overlapping parts of the images are considered further. For each scan line in each image, pixel values are integrated along the scan line, and the resulting function is normalized so as to have a maximum value of one. The root-mean-square difference between each pair of corresponding scan lines is calculated and summed over all such pairs. For stereo pairs this gives a small value, and for random pairs a high value.

Test Image Pairs

As part of the development process, a number of synthetic pairs of images were generated, having known spatial relationships and photometric differences. These were intended to be of a very simple design, but they allowed the various algorithms to be tested. One such example is shown in figure 2 and discussed in the results below.

We also tested the procedures described above on a number of test image pairs that had been acquired with a pair of consumer digital cameras that had shutters wired together so that they would perform synchronously. In each case, one image in the pair showed both global and local differences relative to its reference image. These differences either occurred spontaneously because of inconsistencies in the cameras' automatic exposure calculation; by differences in the relative positioning of bright and dark objects relative to the cameras' sensors; by placing neutral density filters over one camera only; or more likely, by the authors intentionally misadjusting white balance and speed settings. An example of one of these pairs is shown in figure 3.

In our laboratory we are mainly interested in the application of stereographic methods to radiographic images. This is becoming increasingly important as Radiology moves toward the digital acquisition of large 3D datasets. Figure 4 presents a stereo pair of x-ray projections of a breast that were acquired with a breast tomosynthesis system. A single breast exam may acquire many such pairs and these kinds of procedures necessitate automated image correction.

All stereo pairs were processed with the above algorithms and a figure-of-merit was calculated before and after the correction.

Results

Figure 2a and 2b represent right- and left-eye projections of a 3D space consisting of a square, circle and triangle at varying distances from the observer. Colors of the image in 2b were

intentionally changed to test our algorithm's ability to make a suitable correction. Note that colors in each object were altered separately so there is no global adjustment that can adequately reduce the differences. Figures 2c and 2d show the result of the segmentation and depth evaluation, and our use of gray values to label the depth of each object. The corrected version of 2b is shown in 2f.

Figure 3a and 3b are the left- and right-eye views of a stereo pair, where 2b exhibits an overall blue cast, and the small yellow boat on the right appears to have been incorrectly recorded. Manually removing the blue cast did not greatly improve the hue of the small boat. However, the algorithm was able to perform both a global color correction and locally improve the boat's color.

In each of these cases, the algorithm was able to decrease color discrepancies between the images, though the corrections were less than what could have been achieved with a manual procedure.

Figure 4a was considered to be a correct left-projection and 4b was the corresponding right-projection. Figure 4e is the corrected version of 4b. Figures 4c and 4f are the scatter plots associated with the uncorrected and corrected images, respectively.

Discussion and Conclusions

At this stage of development, it is not possible to implement the kinds of corrections considered herein in a fully automatic procedure. This is largely due to the inherent ambiguities in identifying corresponding regions between images – a task that is not always solvable by either computers or humans, but is much more difficult for computers. Also, because 3D shape information is reflected in subtle photometric differences in views at slightly different viewing angles, there is a limit as to how much correction is actually desirable. Issues of this type must rest on the expertise of human observers, until a more comprehensive theory of stereo vision provides guidance. But in the end, if photometric information about a scene has been degraded in one image of a stereo pair, there will not be enough information to correct unambiguously the image.

Acknowledgements

This work is sponsored in part by grant CA80836 from the National Cancer Institute, National Institutes of Health, by Planar Systems, Beaverton, OR; by Grant BCTR0600733 (to the University of Pittsburgh) from the Susan B. Komen foundation; and also by the US Army Medical Research Acquisition Center, 820 Chandler Street, Fort Detrick, MD 21702-5014 under Contract DAMD17-02-1-0549 and contract PR043488. The content of the contained information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

References

1. V.S. Nalwa, A Guided Tour Of Computer Vision (Addison-Wesley, Reading, MA, 1993)
2. J. Canny, "A Computational Approach To Edge Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, 8, 679 (1986).
3. I.J. Cox, S.L. Hingorani, S.B. Rao, B.M. Maggs, "A Maximum Likelihood Stereo Algorithm," Computer Vision and Image Understanding. 63, 542 (1996).
4. D. Geiger, B. Ladendorf, A. Yuille, "Occlusions And Binocular Stereo," Intl. J. of Computer Vision, 14, 211 (1995).
5. S.S. Intille, A.F. Bobick, Disparity-Space Images And Large Occlusion Stereo, Proceedings of the Third European Conference on Computer Vision, pg. 179. (1994).
6. S. Birghfield, C. Tomasi, "Depth Discontinuities By Pixel-To-Pixel Stereo," Intl. J. Computer Vision, 35, 269 (1999).

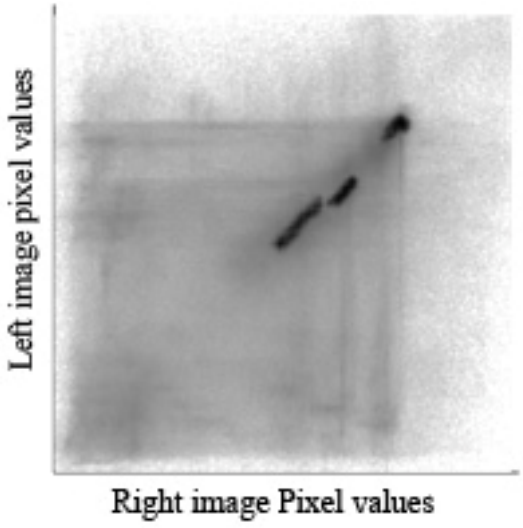
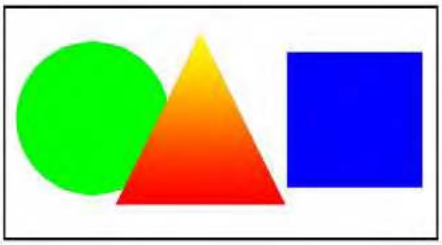
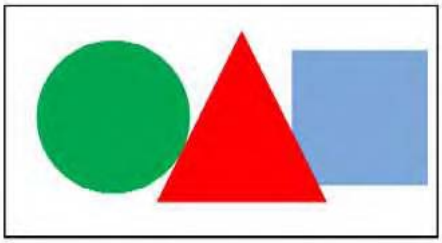
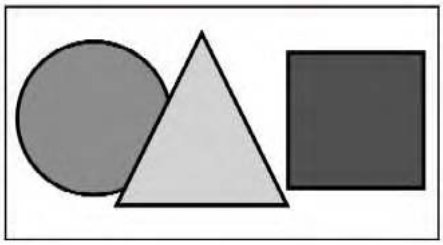


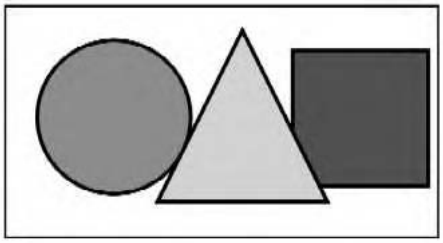
Figure1. Scatter point derived in figure-of-merit calculation for left-eye and corrected right-eye images in figure 3.



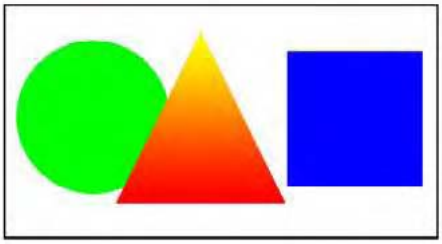


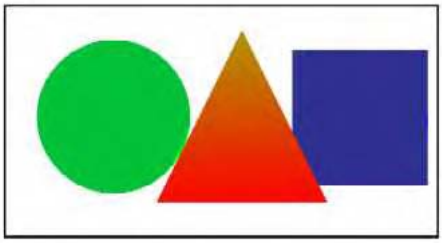
A





C





E

Figure2. A synthetically generated stereo pair of three geometric forms at varying distances from the observer. A and B are the original right- and left-eye images respectively. C and D show how the regions were segmented and labeled with distance (i.e., shade of gray).

A

C

Figure3. A and B represent the left-eye and right-eye images in a stereo pair. Image C is the result of correcting B to match A.

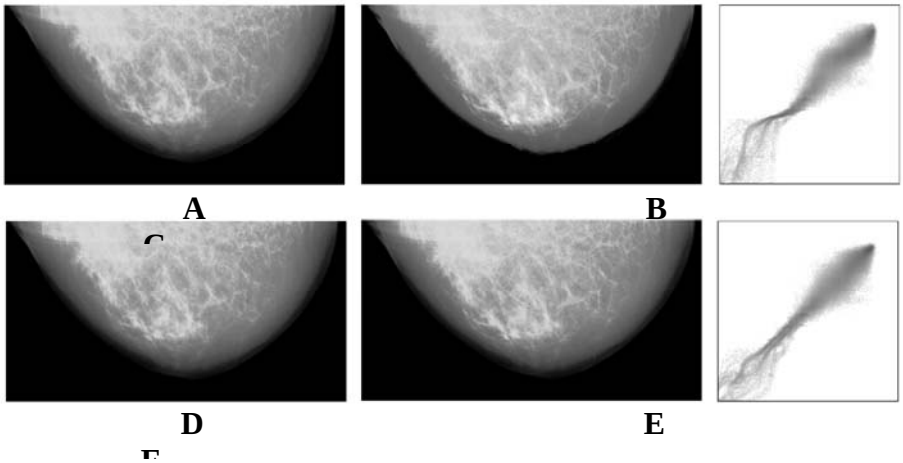


Figure4. A and B are a left- and right-view of breast projections acquired by X-ray tomography. E is the result of correcting B to match A. C and F are the scatter plots generated by the uncorrected and corrected pairs respectively.